

**FIGURE 1**

CGGACGCGTGGGTGCGAGGCCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA  
CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG  
GTTTTCCACCCAGCTTTACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA  
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAACAAGAATTGGGATCCGGCGTGGGA  
GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT  
GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA  
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT  
ATGTCAAGGATAGAATTTCATTCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT  
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG  
GGTGACAATTGGTGTGACCTTTCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC  
CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT  
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCCCTCTTCATCAGAGCTGCATGGTACAC  
AGCTGGCATTGTGGGAGGCCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA  
ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG  
TTTCTTCCACCTACCACCGTGGCTGGTGGCACTCTTTACTCAGTGGCAATGTACGGTGGATT  
AGTTCCTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT  
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT  
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAA**TG**  
**A**AGTGA CT CAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAAATGGGGCAGATATGC  
ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA  
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT  
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT  
TTGGTGAATGTGAAAATAAAGTTTGTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA  
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT  
GCAGAATATTGTAATTAATGTCATAAGTGATTGGAGCTTTGGTAAAGGGACCAGAGAGAAG  
GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA  
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCAATTT  
GCTGAACTTAACAAAACCTGTTCACTCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG  
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC  
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA  
ATAACTTTAAAAC TATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG  
AATACAAACAGTATACTCATG

MLAARLVCLRTLPSRVFHPAFTKASPVVKN SITKNQWLLTPSREYATKTRIGIRRGR TGQEL  
KEAALEPSMEKIFKIDQMGRWV FVAGGA AVGLGALCY YGLGLSNEIGAIEKAVIWPQYVKDRI  
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP  
GPKHLAWLLHSGVMGAVVAPLTI LGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL  
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGVLVLESMFLLYDTQKVIKRAEVSPMYGV  
QKYDPINSMLSIIYMDTLNIFMRVATMLATG GNRKK

**FIGURE 3**

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCGTCTCCGCCTTCTGCAT  
 CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGCGCGTCGTGAGGG  
 GGTCCGCGACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTGGAAG**ATG**TCGG  
 ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCGGCCACCGTC  
 GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCTACCTCTTCCTCTGGCCCGA  
 AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG  
 GTCCAGGAAGTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA  
 CTTGAAACAGGAGCTTTTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG  
 GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTGATCA  
 TGTCACTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTGGTTTGGGA  
 ACACGATTTAAGGCCTGCTATTTACCCCTGGGTATCCTTGGATTCAACTATATCATCGGAGG  
 CTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTCTTAATGTTCA  
 GATACCCAATGGACTTGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG  
 CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC  
 TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC  
 AG**TGA**AGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCCCTCCAGTGCTGGGTG  
 CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC  
 AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAGTTTCACGAT  
 TCTCATTTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAATTGCAAACCTGAC  
 TACATTTTTTGGTGTCTTCTCTTCTCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCTCT  
 AATCTGCTGGCATTGAGCTGGGGCTGGGTACCAAACCTTCCCAAAGGACCTTATCTCTT  
 TCTTGACACATGCTCTCTCCACTTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG  
 CCCATAAAATTTGCTCTGCCCCTTGACAGGTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC  
 ACAACAATCATATTCACGTTATTTTCCCTTTTGGTGGCAGAAGTGTACCAATAGGGGGAG  
 AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATGCTTCGACTGACATCCGTTGTT  
 AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAAGTACCCTGAGTTCATGAGGGCCA  
 CAGATTGTTTATTAATGAGATACGAGGGTGGTGCTGGGTGTTTGTTCCTGAGCTAAGTGA  
 TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC  
 TTTGCGTTTTATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTGCTTTGTGT  
 TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGCTCTGGGCATTG  
 ATTCCCATTTCAATCTCATTTCTGGATATGTGTTTATTGAGTAAAGGAGGAGAGACCCTCATA  
 CGCTATTTTAAATGTCACCTTTTGGCTATCCCCCGTTTTTGGTTCATGTTTCAATTAATTGT  
 GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA  
 AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA  
 TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT  
 TCGTAGGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTGAAGGCCA  
 TGGCTTTTACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT  
 ATTGAGTGGCTGTCACACTTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT  
 TTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT  
 GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG  
 TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA  
 TGTTTTGTTGTTTATCTGTGGCCACAATAAGTTTACTTGTAAATTTTAGAGGCCATTACT  
 CCAATTATGTTGCACGTACACTCATTTGTACAGGCGTGGAGACTCATTTGTATGTATAAGAATA  
 TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCTCTTACCAGTCAGCTGCCTGCGAG  
 CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAAGTTTTTCAAGTTTCAAGGGCAAATGTTT  
 ATGAAGTTATTTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT  
 ATGTTTTCTGGAATAATTTTACCAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA  
 TGACAGTGGATTCTCTTTACAAATGGAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC  
 TTTTGTAAACTAATCCTTTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

**FIGURE 4**

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF  
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLENWICIVITGLAMDMQLLMIP  
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL  
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL  
GDQ

**Transmembrane domain:**

amino acids 98-116, 152-172

**N-myristoylation site.**

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

**Glycosaminoglycan attachment site.**

amino acids 218-222



**FIGURE 5**

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG  
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT  
CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC  
CCTTTGGGGCGGG**ATG**GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG  
CGGGGTTCCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT  
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT  
GGCCTGTGTTCCCCTTGTTTTTGTATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC  
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT  
AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC  
CATTTTGCACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA  
AAAACATTGAAATGCAGCTGCAAGCCATTTCGAATAATTCAAGAGAGAAATGGTGTATTACCT  
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT  
GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA  
AACAGTTATCAGAGGCTAAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG  
AATAATTCCCAAGGGGATGGTGAACATTTTGCACACCCACCCCTCAGAAGTTAAAATGCATTT  
TGCTAATCAGTCAATAGAACCTTTGGGAAGAAAAGTGGAAAGGTCTGAAACTTCCTCCCTCC  
CACAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC  
TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA  
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG  
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG  
CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG**TA**  
**A**TAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAATAAATTATTTAGTC  
CTTACACTG

## **FIGURE 6**

MAAEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP  
LVFDDEEESKLTYTEIHQEYKELVEKLLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP  
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVSDLEHEEMKILREVL  
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS  
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS  
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRLLAEKLKEEVINK

### **N-glycosylation sites.**

amino acids 224-228, 246-250, 285-289

### **N-myristoylation site.**

amino acids 273-279

### **Amidation site.**

amino acids 252-256

### **Cytosolic fatty-acid binding proteins.**

amino acids 78-108

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



GCCTGGTTTTTTGTCTGCAATAGGCGGGCTTAGAGGGAGGGGCTTTTTTCGCCCTATACCTACTG  
TAGCTTCTCCACGTATGGACCCATAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG  
TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTA AAAAAC  
AGTGGA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC  
ATTCTTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT  
CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTG  
GTGAAGCTAGTTTTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAG  
AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCTGCC  
TTCTTTATTTCCCTGGATAACTTGATTGTCTTCTATGTCTGTCTCTATCTTCAACCAGCCATG  
GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAA  
GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTTGTCTATTGTGGCCT  
TGACTGCCGGGACTAAAACCTTACAGCACAACTTGGCAGGAGCTGGATTTCATCACGATGCC  
TTTTTCAGCCCTTCCAATTCCTGCCCTTCTTTTCAGAAGTGAGTGTC CAGAAAAGACAATTG  
TACAGCAAAGGAATGGACTTTTCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC  
ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT  
AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA  
GAACAGCAAACCTCTATTTCTTGGCATTCTGTTAATGGGCTGACTCTGGGCCTTCAGAGGA  
GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTATGGCCACAGTGCATTTTTCAGTAGCC  
CTTATTTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA  
CATGTTCCATTGTCTTGATGGCCAGGTTACCAGTGTCAATTATCACAACAGTGTCTGTCTGG  
TCTTTGCACTTCAGGCCTCCCTGGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA  
TTTTTTTTATAATGCCAGCAAGCCTCAAGTTC CGGAATACGCACCTAGGCAAGAAAGGATCCG  
AGATCTAAGTGGAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA  
CCAAACCCCAAGAGTGATGAGTCAGATGAAGATACTTTCT**TAA**CTGGTACCCACATAGTTTGCA  
GCTCTCTTGAACCTTATTTTTCACATTTTCAGTGTTTGTAATATTTATCTTTTCACTTTGATA  
AACAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT  
CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA  
ATATGAAGAATTATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT  
TTTCTTGGCCTTCAAGCTTCCAAAAAAGCTTGTAATAATCATGTTAGCTATAGCTTGTATAT  
ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT  
CTTCCCTTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGGCCCTAGAGATAGT  
CATTTTGCAAGTAAAGAGCAACGGGACCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC  
CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGTGTAAGAAGCAGT  
TTCTCAGACACAACATCTCAGAATTTTAAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT  
AATAATCTTTTGATGTTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA  
AGTCATTTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT  
GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTATTTTTATTAGTTACT  
AATTCAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC  
CAGATTGTGCTGAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAGGATCACTT  
AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC  
AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTTCATG  
GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA  
AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG  
CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT  
CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTTCATGCAGATGAATATA  
AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG  
GTGGAAATTTGTAATTTAAATAATTATTAAACCT

**FIGURE 9**

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK  
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV  
IFS NFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF  
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI  
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI  
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI  
YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDDTF

**Transmembrane domains:**

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,  
298-318, 328-368

**N-glycosylation sites.**

amino acids 128-132, 204-208, 218-222, 374-378

**Glycosaminoglycan attachment site.**

amino acids 402-406

**N-myristoylation sites.**

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

**FIGURE 10**

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG  
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC  
GTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG  
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA  
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATACATTCCTGCTAGG  
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG  
AAAACAAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT  
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA  
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCCTGCCTTTCTTTATTTCC  
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC  
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA  
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTGTCTATTGTGGCCTTGACTGCCGGGA  
CTAAAACTTTA

**FIGURE 11**

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGGCGGCC  
 GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT  
 CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG  
 AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCCGGCCCA  
 GGTGAAGAAAGAACCGAAAAAGAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG  
 CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCTTCAGATCTAC  
 CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC  
 CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC  
 TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC  
 ATCTGGTTTCGTGCCGACTTCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT  
 CTTTGAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA  
 ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC  
 AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG  
 ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCCTTCACAC  
 AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCAATTGTCTGTATCTATATAATCTG  
 TGCTGTCACTCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTG  
 AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATAACAACTT  
 ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT  
 TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT  
 CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT  
 GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA  
 CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC  
 TACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT  
 GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTACCAAGTTTGCCCTCTGGAGTGTC  
 ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC  
 CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG  
 CTGGGCCCTGCTGCTCTTCAAAATGTACCCCATGATGAGGAGAGGCGGCGGCAGAATAAGAA  
 GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG  
 AGCTGGCTAGCATCCTC**TAG**GGGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG  
 AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA  
 CTGAAGACTCAAGGAGGTGGCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG  
 TGGCCTCCTGCCTCCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA  
 TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC  
 TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTATGTATATGTCTGTGAGCTA  
 TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

MWLRWALS LPPSSCLWAE PGMPSQT PWWASASAN PPGPAWVALCPGSSSPRPWPSLPTSSSG  
SCPTSHTARPIGT CFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL  
GTAIQGQIVGQADTPCFQDFNSSTVASQS ANH THGTTSHRETQKAYLLAAGVIVCIYIICAV  
ILILGVREQREP YEAAQQSEPIAYFRGLRLVM SHGPYIKLITGFLFTSLAFMLVEGNFVL FCT  
YTLGFRNEFQNL LLAIMLSATLTIP IWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI  
ITYAVAVAAGISVAA AFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG  
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNK KAL  
QALRDEASSSGCSETDSTELASIL

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423	2424	2425	2426	2427	2428	2429	2430	2431	2432	2433	2434	2435	2436	2437	2438	2439	2440	2441	2442	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---



**FIGURE 13**

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT  
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA  
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA  
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT  
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC  
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG  
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC  
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT  
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG  
GAACCGAGCCCCAT

**FIGURE 14**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT  
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAA**ATGT**  
GGTGGTTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT  
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT  
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG  
CAGTTTTATGCATTGCTACCATTTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA  
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTGGAACTGAGTTGTTTAGG  
ACTTCTATTGTGGCAAACCTTCAGAAAACAACCCTTTTGGCTGCACATGTAAGTGGAGCTG  
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG  
CAGCCCAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG  
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG  
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT  
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA  
TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG  
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATT**TGAT**GAAAGGAT  
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA  
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA  
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA  
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACTATG

**FIGURE 15**

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI  
AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG  
AVLTFGMGSLYMFVQ TILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG  
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFSGFFLT YIRDFQKISLRVEANLHGLTLYD  
TAPCPINNERTRLLSRDI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 16**

CGGACGCTTGGGCNCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT  
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA  
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT  
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTCTGTTGAGTACA  
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA  
GTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT  
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT  
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

**FIGURE 17**

CCACGCGTCCGCCCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGG  
CCGGGGTGCGGAGCCGACATGCGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC  
CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG  
AGGCTGGAGGCAGGTGCTGTGGTTCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG  
GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCTTGCTCTTCTGCGGCGCCTA  
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT  
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC  
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCCTGATAAAGT  
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT  
TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT  
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT  
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG  
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT  
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA  
CACATTGATCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA  
TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG  
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT  
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT  
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATT  
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAAT  
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC  
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT  
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

## **FIGURE 18**

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRESEVLREYR  
KEHQAYVFLLFCEGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS  
IFGKQLVVSYPDPKVALLRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPVQFF  
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKFSQKHLQ  
LNETSTANHIHSRKDT

### **Important features:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domains:**

amino acids 101-123, 189-211

#### **N-glycosylation sites.**

amino acids 172-176, 250-254

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 240-244, 261-265

#### **N-myristoylation site.**

amino acids 13-19, 104-110, 115-121, 204-210

#### **Amidation site.**

amino acids 27-31

#### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 4-15

#### **Protein splicing proteins.**

amino acids 25-31

#### **Sugar transport proteins.**

amino acids 162-172

18/330

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATT  
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAAC TGCCAGCTCAGAA  
TAGGAAAATAACTTGGGATTTTATATTGGAAGAC **ATG**GATCTTGCTGCCAACGAGATCAGCA  
TTTATGACAACTTTT CAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG  
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC  
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT  
TGCTCACTGCCTACTTTTGTGATTCAACCTTT CAGCCCATTAGCACCTGAGCCAGTGCTTTCT  
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCTTGCCCATTGCCAA  
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC  
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC  
AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA  
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC  
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTTCGCCAAG  
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA  
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCTACTCACCTGCCATTTCCAAAAGATG  
CCTCTTTAAACAAGTGCTCCTTTCTTACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG  
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA  
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG  
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT  
GGAACCGCTTTCTCAGAACTG **TAG**GAAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA  
AAACCAGGTTGAAAGGGGAAAAATAAAAAACAAAACGATGAAACTGCAAAAA

**FIGURE 20**

MDLAANEISIIYDKLSETVDLVRQTHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY  
KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL  
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT  
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFV  
FTHLPFPKDasLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRHCSVAMP  
IEPGDIGYVDTHWKVYVIARGVQPLVICDGTAFSEL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 21**

CCACGGTGTCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG  
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAGTTGTGATC  
AAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT  
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAACTTTCAGAGACTGTTG  
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC  
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT  
TGTGTATAAGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC  
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

CCACACGCGTCCGCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG  
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG  
CCACTGGTGCACGCTGCTAGACCGTGCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC  
CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA  
CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAACC**ATGT**CCCCAGTTCGAAATG  
GACACGTATGCTAAGAGCCACGACCTTATGTACAGTTTCTGGAATGCCTGCTATGACATGCT  
TATGAGCAGTGGGCAGCGGCCAGTGGGAGCGCGCCCAGAGTCGTCGGGCCTTCCAGGAGC  
TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCCTGGAGGGGCTACGCTACACGGCAGTGCTG  
AAGCAGCAGGCAACGCGACACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCT  
CGCCAGCCCATGTGGGGCTGGGCGTGAAGGACACTCCATCCCCGCTGGAAACTGTCCA  
CGCCGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCAC  
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCTGACACCCACCGAGGAGGC  
CTCACTGCCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCGAGTTGCTGCAGG  
AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG  
GATGAGCAGCGTGAGAAGCTGGTGCTGTCGGCCGAGTGCCAGCTGGTGACGGTAGTGGCCGT  
GGTCCCAGGGCTGCTGGAGGTCAACACACAGAATGTATACTTCTACGATGGCAGCACTGAGC  
GCGTGGA AACCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCCAGCTGCGTGAG  
GTCCACCTGCGGCGTTTCAACCTGCGCCGTTACGCACTTGAGCTCTTCTTTATCGATCAGGC  
CAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCCACTCTCATCTCTAGCCAGA  
CTCCGAGACCCCACTGGCCCCATCCCCACCCATACCCAGGTACGGAACCAAGGTGACTCG  
TGGTCTCTGCGCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT  
GCTGCGTGCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACT  
TGATGCAACTCAACACCATTTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTT  
CCCTGGGTCTGCAGGACTACGTGTCCCCAACCTGGACCTCAGCAACCCAGCCGTCTTCCG  
GGACCTGTCTAAGCCCATCGGTGTTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGGAAAGT  
ATGAAAGCTTTGAGGACCCAGCAGGAGCAATTGACAAGTTCCACTATGGCACCCACTACTCC  
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTCACTCCCTGCACGTCCA  
GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC  
AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTTCTTACTTTCTCT  
GACTTCTTGGAACACGAGACCGTTTTGACCTGGCTGTCTCCAGCTGACCAACGAGAAGGT  
AGGCGATGTGGTGACTACCCCGTGGGCCAGCTCTCTGAGGACTTCATCCAGCAGCACCGCC  
AGGCTCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC  
TACAAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA  
GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAAACGGAAGGCTCTGGAGGGCATT  
TCACTCAACTTTGGGCAGACTCCCTGTGACTGCTGAAGGAGCCACATCCAACCTCGGCTCTCA  
GCTGAGGAAGCAGCCCATCGCCTTGACACGCTGGACACTAACTACCTAGCATCTTCCAGCA  
CCTGGACGAACTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA  
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC  
ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTG  
GAGTGGACAAGCACTGGCAGTGGCCCCGATGGAAGCTGCTATTACGCGGTGGCCACTGGG  
ATGGCAGCCTGCGGGTGACTGCACCTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGTGCCAC  
CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG  
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGATAGGCCTGGCAC  
CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT  
GAACCTGACATGGCTGTGTCTGAGTCTGAGGATGGAAGCTGATCATACACACTGTATCGCCG  
CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC  
TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCC  
CAGGTACCTACTCCTTGACCTGTATTAGTCAATGGGAAGTTGCGGGGCTTCACTGCCCCCT  
GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCAGT  
GCGCCCTGCACATCTCCAACATAACACACTGCTCCCCGGCGCCTCCCTTGCCCATGAAG  
GTGGCCATCCGACAGCTGGCCGTGACCAAGGAGCGCAGCCAGTGTGGTGGGCCTGGAGGA  
TGGCAAGCTCATCGTGGTGGTTCGCGGGGCGAGCCCTCTGAGGTGCGCAGCAGCCAGTTTCGCGC  
GGAAGCTGTGGCGGTCTCTCGCGGCGCATCTCCAGGTGTCTCGGGAGAGACGGAATACAAC  
CCTACTGAGGCGCGCT**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCCGCCCCGGCAGGCCCTG  
GCCCGGGAGGCCCCGCCCCAGAAGTCGGCGGGAACACCCCGGGTGGGCAGCCAGGGGGTGA  
CGGGGGCCACCTTGCCAGCTCAGGGATTGCGGGCGCATGTTACCCCTCAGGGATTGGCG  
GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCGGCCCTGAGGGCCAGCACTGGCGTCT

**FIGURE 23**

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL  
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN  
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAEELETP  
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP  
 LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTTPVSSPSQTPRPQPGPIPPHTQV  
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL  
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSPKIGVVPKHAQLVREKYESFEDPAGTIDKFH  
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAQARLESPADVKEIIP  
 EFFYFPDFLENQNGFDLGLCLQTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW  
 IDLI FG YKQRGPAEEALNVFYCYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP  
 HPTRLSAEEAAHRLARLDTNPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF  
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGSLRVTALPRGKLL  
 SQLSCHLDVVTCLALDTCGIYILISGSRDTTCMVWRL LHQGGLSVGLAPKPVQVLYGHGA AVS  
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPFIHFLALGSEGQIVVQSSA  
 WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALVTEDFVLLGTAQCALHILQLNTLLPAA  
 PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS  
 GETEYNPTEAR

**N-glycosylation site.**

amino acids 677-681

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 985-989

**Tyrosine kinase phosphorylation site.**

amino acids 56-65, 367-376, 543-551

**N-myristoylation site.**

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,  
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,  
 873-879, 912-918, 954-960

**FIGURE 24**

**CGG**ACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC  
CACGGCCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGGCTACTCAT  
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTTGGGGCTCTTC  
TGGACCCTTAAGTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTGCCTCCTT  
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC  
GCACACTCCGTTACCACACTGGGTCATTGGCATTGTGGAGCCCTCATCCTGACCCTTGTGCAG  
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC  
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC  
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA  
AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTTCGTCTGACAAAGTCAAGAG  
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCTGTCTTCTTTT  
TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCACCTCAACTATTAC  
TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT  
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG  
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC  
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGGCCCTGATCCAGGACTGC  
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT  
AAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG  
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC  
GTCTCTATTAAAAATACAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC  
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA  
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA  
AAAGATTTTATTAAAGATATTTTGTTAAGTCT

**FIGURE 25**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLGF  
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ  
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK  
NAFMLLMRNIVRVVLDKVTDLLLFFGKLLVVGGVGVLSEFFFFSGRIPGLGKDFKSPHLNYY  
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYSKSLKILGKKK  
EAPPDNKKRKK

**FIGURE 26**

GAGTCTTGACCGCCGCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCGCT  
 GGCT**ATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC  
 TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTT  
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC  
 ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG  
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG  
 CCAGTCAATGTTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAACAAGATGA  
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT  
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA  
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT  
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG  
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC  
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTGCAGCG  
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA  
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC  
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA  
 GCGGCTCCAGGAGTTCCCTTGAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC  
 AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA  
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCATTTTGGGTTCAAGCACAAGTT  
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT  
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG  
 TACCATGGCCTGGAACCTGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC  
 CTTTGCACCAACCTCGTCATCTCCCAGGGGCTTTCCCTGTACTGCTCTCTCATGGAGGGCAC  
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA  
 AGTCCTTTGTGTGTTTCGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCTGGTGATGGCTGCC  
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC  
 GGACAGGAAGAAGCTTTTTTGGGAGGGCGTTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA  
 TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT  
 CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAATTTGATTCTTCCAGAATGACCTTCTTATT  
 TATGTAAGTGGCTTTTCAATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT  
 TTTTATTAAATAAAATGCTTATTTTAGGAAA

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

**FIGURE 27**

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF  
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD  
LEVPAIEDIFRDEEEDEEHSGNDSGSEPSEKTRLEEEIVEQTMRRRQRREWEARRRDILF  
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH  
VSRHNRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR  
LQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFL  
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL  
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP  
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNFHFDLSVIELKAEDRSKFL  
DALISLLS

**FIGURE 28**

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA  
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT  
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT  
GGTTCAGTTTCTGGGTGGCAAGAACTTGAAACTGCATTTCTTGAGCATAAAGAACAGTTTC  
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT  
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA  
CGATACCC

0909724 412004  
100214 12/08/60



**FIGURE 29**

## **FIGURE 30**

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH  
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPGQGEPPGQSGKGDKGEMGSPGAPCQKRF  
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET  
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT  
FSGHLIKAEDD

**Important features:**

**Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 72-75

**C1q domain proteins.**

amino acids 144-178, 78-111 and 84-117

ACTCGAACGACGAGTGGCTTCGCGGACCCAGGACCCCTCGGGCCGACGACGAAAGACTG  
AGGCCGCGGCCTGCCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCGGACAGAAGATG  
CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGACAG  
GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG  
ACCACGGTGGCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT  
CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC  
AGAACCAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG  
GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCC  
CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC  
TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCGCCTGCGC  
CTGCCCCGCCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT  
CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG  
AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACAGCTGGAG  
CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC  
CCGCATTGCCCAGCTGCGGCCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG  
TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG  
CTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGCCCCCTG  
GGTGC GCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCGCCCCA  
AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACC  
ACCACAGCCACAGTGCCCAACCACGAGGCCCGTGGTGC GGGAGCCACAGCCTTGTCTTCTAG  
CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTCCA  
CTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTC  
AATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT  
CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA  
CGCCGAGGGCCACCACGCTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC  
GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA  
TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG  
AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGG  
CCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGGCCATACACCCCCAGCCGTCCA  
CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCCG  
CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG  
GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTC  
GGAACCTGGAGGGAGTGAAGGTCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG  
AGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC  
CAGTCACCCCTCCACGCAAAGCCCTACATCTAAAGCCAGAGAGAGACAGGGCAGCTGGGGCCG  
GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC  
CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA  
CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCAGAAC  
CGAGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG  
GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA  
CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG  
TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA  
GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTTATTCTG  
GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATG  
AAGGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

**FIGURE 32**

MCSRVPLLLLPLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDPVPPDTVGLYVFEN  
GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR  
RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP  
GILDTANVEALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAG  
NTRIAQLRPEDLAGLAALQELDVSNSLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG  
PWVRESHVTLASPEETRCHFPKPNAGRLLLELDYADFGCPATTTTATVPTRPVVREPTALS  
SSLAPTWLSPTAPATEAPSPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE  
GFTGLYCESQMGQTRPSPTPVTPRPPRSLTLGIEPVSPTSRLRVGLQRYLQGSSVQLRSLRL  
TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPA  
VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRGRAMAAAAQDKGQVGPGAG  
PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

GAATCATCCACGCACCTGCGAGCTCTGCTGAGAGAGTGC AAGCCGTGGGGGTTTTGAGCTCAT  
CTTCATCATT CATATGAGGAAATAAGTGGTAAATCCTTGAAATACAATGAGACTCATCAG  
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG  
AAGAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG  
ACCCCGCCCAACAGCAGACTGGATTATCTCTATAACCTCCTTTTTCAACTCCAGAGTTTCA  
TTTTCACTTGTCTCCAACTGAGAGTTTGTATTCTATGCCATAACAGAATTC AACAGCTGG  
ATCTCAAAACCTTTGAATTCACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG  
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT  
TGACACCATTGCCTATCTGTGAGGAAGCTGGCAACATGTACACCTGGAATCCTAGGTTTGA  
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTTGCTCATCTGCATCTAAATACTGTC  
TTCTTAGGATTGAGAACTCTTCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC  
AAAACCTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCGTGATGGAATCA  
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA  
ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA  
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTGGCATACATCAGTGGAACTCT  
TTGAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC  
TCAAATACTGTAATGAGAACTATAAAATTTGGAGCATGTACATTT CAGAGTGTTTTACATTCA  
ACAGGATAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG  
CACAAATGCCACACATGCTTTTCCCGAAATTAATCTACGAAATTCCAATATTTAAATTTTGGC  
AATAATATCTTAACAGACGAGTTGTTTAAAGAACACTCAACTGCCTCACTTTGAAAACCTCT  
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC  
CCTTGGAACACTTGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTGCTCA  
TGGCCAGAACTGTGGTCAATATGAATCTGT CATACAATAAATTGTCTGATTCTGTCTTCAG  
GTGCTTGCCCCAAAGTATTCAAATACTTGACCTAAATAATAACCAAATCCAAACTGTACCTA  
AAGAGACTATTCTATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTCTAATGAT  
CTCCCTGGATGCAGTCATTTCTAGTAGACTTTCACTTCTGAACATTGAAATGAACTTCATTCT  
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA  
ATCCATTCCGGTGTACCTGTGAATTAAAAAATTTCAATTCAGCTTGAACATATTCAGAGGTC  
ATGATGGTTGGATGGTCAGATT CATAACCTGTGAATACCCTTTAAACCTAAGGGGAACCTAG  
GTTAAAGACGTTTCATCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTCAACATTG  
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCC  
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA  
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTATTTTCATACAGTGAACATGATTCTC  
TGTGGGTGAAGAATGAATTTGCAATCCCCAATCAGAGAAGGAAGATGGTTCTATCTTGATTGC  
CTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA  
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT  
ATGAATTCTACTTTGCCACCACAATCTCTTCCATGAAAATTTCTGATCATATAAATCTTATC  
TTACTGGAACCCATTCCATTCTATTGCATTTCCACCAGGTATCATAAACTGAAAGCTCTCTCT  
GGAAAAAAGCATACTTTGGAATGGCCCAAGGATAGGCGTAATGTGGGCTTTCTGGGCCAA  
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACGAGAGAAATGTATGAACCTGCAGACA  
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT  
ATAAATCCCACAGTCTTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA  
CAACCTTTATGATGGCAATTTGACAATATTTATTAAAAATAAAAAATGGTTATTTCCCTTCATA  
TCAGTTTCTAGAAGGATTTCTATAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG  
GCTTATGGA AAAAGGTGTTCTATCCAGGATTGTTTATAATCATGAAAATGTGGCCAGGTGC  
AGTGGCTCACTTTGTATATCCAGCAGCATATGGGAGGCCAAGGTGGGTGATCCACGAGGTCAA  
GAGATGGAGACCATCCTGGCCAACATGGTGA AACCTGTCTCTACTAAAAATACAAAATTA  
GCTGGGCGTGATGGTGCACGCCTGTAGTCCAGCTACTTGGGAGGCTGAGGCAGGAGGAATCG  
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACCTCCAGCTGGT  
GACAGAGCGAGACTCCATCTCAAAAAAGAAAAAAGAAAAAATGGAAAACATCC  
TCATGGCCACAAAATAAGGTCTAATTCAATAAATTTATAGTACATTAATGTAATATAATATTA  
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGAAAAAACATATTAATAT  
GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAA  
ATAAAAGTGTAAAGAAATCTATACAGATGTAGTAACAGTGGTTTGGGCTGGGAGGTTGGA  
TTACAGGGAGCATTTTGATTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTAGAAATGA  
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

**FIGURE 34**

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ  
LQSSDFHSVSKLRVLILCHNRIQQDLKTFFFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL  
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP  
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL  
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR  
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNfANNILtDELfKRTIQLP  
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNdENCsWPETVVNMNLSYNKLS  
DSVFRCLPKSIQILDlnnnQIQTVPKETIHLMALRELNIAFNFLTDLPgCSHFsrLSVLNIE  
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN  
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFcCLHFDLPWYLRMLGQCTQTWHRV  
RKTTQEQlKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI  
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHhNLFHENSdHIILILLEPIPFYCIPTRYHK  
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTfTELNEESRGSTISLM  
RTDCL

**FIGURE 35**

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTGCGCGGGAAAGG  
 GAGGGAAGAAGGAAGGGCGGGGCGGGCCCCCTGCGCGCGCGCCCTGCGCGCGCCCTGTCCGCCCCGGC  
 CCAGCCCAGCCCAGCCCCGCGGGCCGGTACACGCGCAGCCAGCCGGCCGCTCCCGCGCCCAAGCGCGCCGCT  
 CTGCTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCGCCGGCGCCCCGGTGACCGTGA  
 CCCTGCCCTGGGCGCGGGGCGGAGCAGGC**ATGT**CCCCGCGGGGACCGCTACCCAGCGCTGGCCCTGGTGCTC  
 CTGGCAGTGACCTGGCCGGGGTTCGGAGCCCAGGGCGCAGCCCTCGAGGACCTGATTATTACGGGCAGGAGAT  
 CTGGAGCCGGGAGCCCTACTACGCGCGCCCGGAGCCCCGAGCTCGAGACCTTCTCTCCGCGCTGCCTGCGGGGC  
 CCGGGGAGGAGTGGGAGCGCGCCCCGAGGAGCCAGGCCGCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCC  
 AAGAGGGAGAAGTCGGCTCCGGAGCCGCTCCACCAGGTAAACACAGCAACAAAAAAGTTATGAGAACCAGAG  
 CTCTGAGAAGGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTC  
 TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA  
 CATCGAGGGAGACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTATGACGGAGCGTGGTGCGCGGGAAG  
 AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCTGACCAGATTCACTGGTGTCACTCAAGGGA  
 GGAACCTCCCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACAGTGGGT  
 ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT  
 ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGA  
 GAATGGAGATCCTGGGCTGCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT  
 GATGACCTGGATTTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCC  
 CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC  
 ACCCTGGGGAGCATGAAGTCGGTGAGCCCCGAGTTCCACTACATCGCGGGGGCCACGGCAATGAGGTGCTGGGC  
 CGGGAGCTGCTGCTGCTGCTGGTGCACTTCCGTGTGTGAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT  
 GGTGGAGGAGACGCGGATTACGTCCTCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCT  
 CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACCTTCCCTGATTTA  
 AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC  
 TGAGTGGTTTTCTGTCGGAATAATGCCACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAAATCC  
 CTTTTGTGCTGGGCGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCCC  
 TGGAAAGACGCAGGAACACACCCCCCACCCTCGATGACCACGTGTTCCGCTGGCTGGCCTACTCCTATGCCTCCAC  
 ACACCGCTCATGACAGACGCCCCGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA  
 ATGGGGCCTCCTGGCACACCGTCGCTGGAAGTCTGAACGATTTAGCTACCTTCATACAACTGCTTCGAACTG  
 TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC  
 TCTGATCGTGTTCATGGAGCAGGTTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCC  
 CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCTC  
 CTGAACCTGGAGAGTATGTGGTCACAGCAAGGCCGAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG  
 CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGG  
 AGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT  
 GGG**TGA**CCCTCCTGGGCCCTTGAGACTCGTCTGGGACCCATGCAAATTAACCAACCTGGTAGTAGCTCCATAG  
 TGGACTCACTCACTGTTGTTTCTCTGTAATTCAGAAGTGCCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCC  
 CAAAAGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTCTTTGTTCCATTATCCAAATAACTTGGACAGAGCA  
 GCAGAGAAAAGCTGATGGGAGTGAGAGAATCAGCAAGCCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGG  
 GAGCCTGTCCGTTAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCTGTTTGGTGCGTGCCAGCAAGG  
 GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCA  
 TTTGAGATGCTCCAGGCGTCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA  
 ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT  
 TTTTCAACAAAGGAGTGTGTTTCAGAAAAGGAGAGAGAGGCTGAGATCATTCAGGAGTTTGTGGGCGAGCAAGCA  
 TGGAGCTTCTTGACAAATTTCTGGTCCATAAAACACCCCAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT  
 CCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC  
 CTGCTAGGACTGGAAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTCTCTCTGCTTGAAGTATTGCCCCCTGTGTG  
 GAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA  
 GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT  
 CTTGCAAGAAAAAAGTATGTCTCACTTTTTGTTAATGTTGCTGCCTCATTGACCTGGGAAAAAATGAAAAA  
 AATAAAGCAAATGGTAAGACCCTTAAA

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

**FIGURE 36**

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPLP  
AGPGEEWERRPQEPRPPKRA TKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS  
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA  
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF  
EGNSEKEIPVLNELPVP MVARYIRINPQSWFDNGSICMRMEILGCPLDPNNYYHRRNEMTT  
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF  
HYIAGA HGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNP DGYEKAYEGG  
SELGGWSLGRWTHD GIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA  
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYS YAST  
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFS YLHTNCFELSIYVGCDKYPHES  
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL  
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTL SKTNMARI REIMEKFGKQPVSLPARR  
LKL RGRKRRQRG



1990-1991		1991-1992		1992-1993		1993-1994		1994-1995		1995-1996		1996-1997		1997-1998		1998-1999		1999-2000		2000-2001		2001-2002		2002-2003		2003-2004		2004-2005		2005-2006		2006-2007		2007-2008		2008-2009		2009-2010		2010-2011		2011-2012		2012-2013		2013-2014		2014-2015		2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021		2021-2022		2022-2023		2023-2024		2024-2025		2025-2026		2026-2027		2027-2028		2028-2029		2029-2030		2030-2031		2031-2032		2032-2033		2033-2034		2034-2035		2035-2036		2036-2037		2037-2038		2038-2039		2039-2040		2040-2041		2041-2042		2042-2043		2043-2044		2044-2045		2045-2046		2046-2047		2047-2048		2048-2049		2049-2050		2050-2051		2051-2052		2052-2053		2053-2054		2054-2055		2055-2056		2056-2057		2057-2058		2058-2059		2059-2060		2060-2061		2061-2062		2062-2063		2063-2064		2064-2065		2065-2066		2066-2067		2067-2068		2068-2069		2069-2070		2070-2071		2071-2072		2072-2073		2073-2074		2074-2075		2075-2076		2076-2077		2077-2078		2078-2079		2079-2080		2080-2081		2081-2082		2082-2083		2083-2084		2084-2085		2085-2086		2086-2087		2087-2088		2088-2089		2089-2090		2090-2091		2091-2092		2092-2093		2093-2094		2094-2095		2095-2096		2096-2097		2097-2098		2098-2099		2099-2100		2100-2101		2101-2102		2102-2103		2103-2104		2104-2105		2105-2106		2106-2107		2107-2108		2108-2109		2109-2110		2110-2111		2111-2112		2112-2113		2113-2114		2114-2115		2115-2116		2116-2117		2117-2118		2118-2119		2119-2120		2120-2121		2121-2122		2122-2123		2123-2124		2124-2125		2125-2126		2126-2127		2127-2128		2128-2129		2129-2130		2130-2131		2131-2132		2132-2133		2133-2134		2134-2135		2135-2136		2136-2137		2137-2138		2138-2139		2139-2140		2140-2141		2141-2142		2142-2143		2143-2144		2144-2145		2145-2146		2146-2147		2147-2148		2148-2149		2149-2150		2150-2151		2151-2152		2152-2153		2153-2154		2154-2155		2155-2156		2156-2157		2157-2158		2158-2159		2159-2160		2160-2161		2161-2162		2162-2163		2163-2164		2164-2165		2165-2166		2166-2167		2167-2168		2168-2169		2169-2170		2170-2171		2171-2172		2172-2173		2173-2174		2174-2175		2175-2176		2176-2177		2177-2178		2178-2179		2179-2180		2180-2181		2181-2182		2182-2183		2183-2184		2184-2185		2185-2186		2186-2187		2187-2188		2188-2189		2189-2190		2190-2191		2191-2192		2192-2193		2193-2194		2194-2195		2195-2196		2196-2197		2197-2198		2198-2199		2199-2200		2200-2201		2201-2202		2202-2203		2203-2204		2204-2205		2205-2206		2206-2207		2207-2208		2208-2209		2209-2210		2210-2211		2211-2212		2212-2213		2213-2214		2214-2215		2215-2216		2216-2217	
-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--	-----------	--

MRPGLSFLALLFFFLGQAAGDLGDVGPPIPSPGFSSFPFVDSSSSSFSSSSSRSGSSSSSRSLGS  
GGSVSQLFSNFTGSVDDRGTQCQSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV  
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ  
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH  
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYSPQHHPNKGlyWVAPLNTDGRLLLEYRlyNTLD  
DLLLYINARELRITYGQSGTAVYNNNMYVNMYNNTGNIARVNLTNTTIAVTQTLpNAAYNNR  
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDDTLQVLNTWYTKQYKPSASNAF  
MVCGLYATRMTNTRTEEIFYYYDTNTGKEGKLdivMHKMqEKVQSINYNPFdqKLYVYNDG  
YLLNYDLSVLQKPO

**FIGURE 39**

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC  
CCTCCTCCCCTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT  
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC  
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG  
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT  
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA  
TGTACAACACCGGNATATTGCCAGAGTTAACCTGACC

**FIGURE 40**

TCTCGCAGATAGTAAATATCTCTCGAAAGGCGAGAGAAAGCTGTCTCCATCTTGTCTGTAT  
CCGCTGCTCTTGTGACGTTGTGGAG**ATG**GGGAGCGTCTCGGGGCTGTGCTCCATGGCGAGCT  
GGATACCATGTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCTTAGTGGAAC  
AACTCCACTGTAACCTAGATTGATCTATGCACATTTTCTTGCTTGTGAGATTTGTAGCTTG  
TGTAATGTTGATACCGGAATGGAAGAACAACTGAATAAGATTCTGGATTTGTGAGAATG  
AGAAAGGTGTTGTCCCTGTGAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT  
GGTTTGGCTATGTTCTATCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA  
TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA  
TTATTGGGGCATTCTTTCATTCCAGAAGGAACTTTTACAACCTGTGTGGTTTTATGTAGGCAT  
GCAGGTGCCCTTTTGTTCATCCTCATACAACCTAGTCTTACTTATTGATTTTGCACATTCATG  
GAATGAATCGTGGGTTGAAAAATGGAAGAGGGAAGCTCGAGATGTTGGTATGCGAGCCTTGT  
TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCTGTTCTTTGTCTAC  
TACACTCATCCAGCCAGTTGTTTCAAGAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTG  
CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT  
TGTTACAGTCTTTCAGTAATTACAGTCTACACAATGTATTGTACATGGTCAGCTATGACCAAT  
GAACCAGAAACAAATTTGCAACCCCAAGTCTACTAAGCATAAATTGGCTACAATACAACAAGCAC  
TGTCCTCCAAAGGAAGGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC  
TCTTTTTTGTGTGTGATTTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA  
CTGACTCTAACAAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC  
ACTGGAGGATGGGGACGATGTTTCAACGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA  
GTTATTCCTTCTTTCACTTCATGCTTTTCTGGCTTTCACCTTTATATCATGATGACCCCTTACC  
AAGTGGTCCAGGTATGAACCCCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAA  
AATCTCTTCCAGTTGGATTGGCATCGTGTGTATGTTTGGACACTCGTGGCACCACCTTGTTT  
TTACAAATCGTGATTTTTGACT**AG**AGTGAGACTTCTAGCATGAAAGTCCCACCTTTGATTATTGC  
TTATTTGAAAACAGTATTCACCACTTTTGTAAAGTTGTGTATGTTTTGTCTTCCCATTGAAC  
TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTGATTTTGTATTTTCTTACCAA  
GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT  
TAGTAAAAGTGGCCATTATTGGGCTTATTCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA  
ACAAATTTGTTTGACTATTTTTAAATTTATATTAGACCTTAAGCTGTTTTAGCAAGCATTTAA  
GCAAATGTATGGGTGCCTTTTTGAATATTTGATGTGTTGCCCTGGCAGGATACTGCAAGAAC  
ATGGTTTATTTTAAATTTTATAACAAGTCACTTTAAATGCCAGTTGTCTGAAAAATCTTATA  
AGGTTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG  
TTATGGATGGAGGTGTGCGTACTAAATTGAATAACGAGTAAATAATCTTACTTTGGGTAGAGA  
TGGCCTTTGCCAACAAAGTGAACCTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTTCA  
GGAAATGTTTGAACCTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTTAGA  
AGGAAGTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCACGGTAGCTCACCCCTT  
GGTAATCCCAGCACTTTGGGAGCTTAAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA  
GCTTGGCACATGGTGAACCTGTTCTATAAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC  
CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAGAGCCAAAGTTGCACTGAGCAAGTCA  
CGTCACTGCACCTAGCTGGCAGACAGTAAGCCAAAAAATATATATATTGAAATCAAGG  
AGGCAAAATTTTGACAGGGAAGGAAGTAACTGCAAAACCCTAGGCTTTAGTAGGTACTTAT  
ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAAATGAAGACACTGAAATACAGACTTAAATA  
GCTCAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAA  
AAATAATTTCTATTCAAAATACATGCATATTGATTTACACCTTCATAGTGATAATTAATGT  
GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT  
AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAGATGTTTAAATGCTTTTGTAT  
TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT  
ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAATCTACAAT  
TCTTCTTTAAAAATATTTTAAATGTGAATGGAATATAACAATTCAGCTTAATTTCCCAACC  
TTATTTCTGTGTAGACATTTGATTTCCAAATTTTGAATGGCTGTGTTTTACCTCTAAATAA  
ATGAATTCAGAGAAAAA

**FIGURE 41**

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME  
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG  
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM  
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI  
LPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV  
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH  
RAVDNERDGVTYSSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSSWIGI  
VLYVWTLVAPLVLTNRDFD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 42**

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT  
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC  
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAAGTAGATTGATCTA  
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG  
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT  
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT  
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT  
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC  
 CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC  
 ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT  
 GTTGATAACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG  
 GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG  
 GCTANGTTCTATNTTCTTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG  
 AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG  
 GGGC

**FIGURE 44**

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC  
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATAACCATGTTTGTGTGGAAGTGCCCCGTGTT  
TGCTATGCCGATGCTGTCCTAGTGGAACAACCTCCACTGTAAC TAGATTGATCTATGCACTT  
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC  
GAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG  
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTA  
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT  
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC



**FIGURE 45**

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTTGGATTGATCTATGCACTTTTTCTTG  
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA  
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC  
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTACT  
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT  
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACCTTT  
ACAACGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACAACTAGT  
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGA  
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA  
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC  
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

**FIGURE 46**

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTCGCGGGCCGGCGCCGGCCTCTCCAAT  
 GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTCGAGACCGGGGCGAG  
 TCCTGTGAAAGCAGATAAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCCGGCGGGGCTGTGCG  
 ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC  
 GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC  
 CCTGTTTGTGTGTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC  
 TGGGCGAATCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTGAATCTGCGAGTG  
 AAGAGGGACGAGGGAAGAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAGAAAGACCCAGAT  
 CAGCAAAAAAGAAAGATGGGGCCCCCGAGCCTCGTGTGCTGCTTGGCTGTCCGCAACTGTGTTCTCCCTGCTGGG  
 TGGAAGCTCGGCCTTCCCTGTGCGACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCA  
 ACATCATCTTGGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC  
 ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGTGCCCTCACGCTCCTC  
 CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC  
 AGGCACAGCAGGAGGCCGACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG  
 TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGCGACTCCTTAAAAACTCCCG  
 CTTTTATAACTACACGCTGTGTCCGAACGGGGTGAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA  
 CAGACCTCATACCAATGACAGCGTGAGCTTCTTCGCGACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTCT  
 ATGGTCATCAGCCATGCAGCCCCCAGGGCCCTGAGGATTACAGCCCCACAATATTCACGCCTCTTCCCAAACGC  
 ATCTCAGCACATACGCCGAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC  
 CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGTGGAC  
 GACTCCATGGAGACGATTTACAACATGCTGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC  
 CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG  
 TCCCGTTTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG  
 GCCCCACCATCCTGGACATTGACAGGCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT  
 GGACACGGAGCGGCCGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG  
 AGAGAGGCAAGCTGTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC  
 CAGCGTGTGAAGGACCTGTGTGTCAGCGTGTGAGTACCAAGCGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG  
 TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC  
 TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC  
 AGCCTGGCCGGACGCCGGAAGAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCGAGTCGCTCCATCCG  
 CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA  
 CCAAGCGGCACTGGCCAGGGGCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCACTGGCACTGGAGGC  
 CTTCCCGACTACTCAGCCGCCAACCCTATTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA  
 GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA  
 CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCCAGAAGATGTGAC  
 TGTACAAAATCAGTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTTCAG  
 GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC  
 TCAAGCGCCTGCAGAACACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACAGCACTGG  
 CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCTGACACAGCGCCAACAATAACAGTACTGGTGCAT  
 GAGGACCATCAATGAGACTCACAAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCA  
 ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG  
 CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGATGGAGG  
 AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG  
 GACAACCTGTGGGAAGGCTGGGAAGGTTAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA  
 CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAAGC  
 AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC  
 ATTTTTGCCCCCTGCTTTTGTGTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTCGTATCAAAAAGTC  
 ACCACTAACCCCTCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGAGATTTCTTGGAAATTTCT  
 TCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCAGTCTTCTAAATCCTCTTATTCTT  
 TTGGTTTGTACAAAAGAAGGAACATAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG  
 TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCCTCTGAAGAAA  
 CTGCCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAAT  
 CCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAGAAAAA



**FIGURE 48**

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA  
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAG**ATG**CTGGGCCTCCTGGGGAGCACAGCCC  
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTCTGCTGCTGCTGCTGCTGCTGGCCACC  
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA  
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC  
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC  
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA  
CGCTCGCT**TGA**GGCTGCTGTGCGCCGGTGCCTGTGGACAGCAGCTGCCCCCTGCCCTCCCATCTG  
TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG  
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTAATATGGC  
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGAGAGCAAGG  
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTCCCAGTGAGCCCCAGAAATG  
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC  
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT  
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT  
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

**FIGURE 49**

MLGLLGSTALVGWITGAAVAVLLLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRRAQWPFR  
RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHHPRHAR

**FIGURE 50**

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTA  
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG  
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCA  
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC  
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC  
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC  
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC  
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTTCTGTCCATCTGGCTGGCTACCCGCCG  
TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG  
AGATCTACCAGGAAGACCAGATCCATTTTATGTGCCCACTGGCACGGCAGGGAGACTTCTAT  
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGGCTTGTGGAGGCCATTGACACCCA  
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC  
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT  
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA  
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC  
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAGTAACCC  
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACCTGAGCAGACTCTCCAGCAGACTCT  
CCAGCCCTCTTCCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC  
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA  
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTTTCAGACTC  
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAA  
AAAAA



**FIGURE 52**

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT  
GCCCCGCGCCAGTCCATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT  
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA  
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA  
GACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT  
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA  
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTAT  
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT  
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAG  
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT  
AGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAATA  
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT



**FIGURE 53**

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGLTLHI  
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF  
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS  
KKKLKEEKRNKSKKK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 54**

CCCGGGAACGTGTTTCCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC  
CCCGCGCCCAGTCATGACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC  
TGCTCAGTGCGGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC  
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA  
CACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA  
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT  
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG  
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA  
TTGCACTAATCCGAGCCAACACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG  
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA  
CCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA  
AATAATAAATTTTAAAAAACTTA

**FIGURE 55**

CCGAAAGTCCCGTCCGGACCCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC  
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG  
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA  
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT  
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT  
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG  
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC  
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAA  
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

**FIGURE 56**

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTCTTGGCTAAAATCGGGGGAG  
TGAGGCGGGCCGGCGCGGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG  
ACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG  
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT  
TATCATAGATGCAGCTGTTATTTATCCACCATGAAAGATTTCAACCACTCATACCATGCCT  
GTGGTGTATATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA  
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG  
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG  
CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT  
TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTGAACACATCTGAT  
TTCCACAGCACAACAGCCCTGCATGGGTTTGTTTGTTTTTTTACTGCTCACTCCCAACCTT  
TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT  
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT  
ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT  
TGTA AAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA  
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT  
CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTTAGTTTTTTAAAATATTCCGTGG  
TCAA AATTCTTCCTCACTATAATTGGTATTTACTTTTACC AAAAATTCTGTGAACATGTAAT  
GTA ACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT  
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT  
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC  
CACATCCACCACTG

57/330

**FIGURE 57**

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDAAVIYPTMKDFNHSYHACGVI  
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK  
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 58**

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGCGGCGCGACACCGGGCTCCGGAACC  
ACTGCACGACGGGGCTGGACTGACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG  
CTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC  
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT  
TTCAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGC  
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG  
CTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG  
ATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT  
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTAAGTTTGGC

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTTCAGAATGCATGAC  
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG  
GTGGATTATCATAGATGCAGNTGTTATTTATCCACCATGAAAGATTTCAACCANTCATACC  
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA  
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT  
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT  
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC  
ATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

**FIGURE 60**

GGACACCGGGTTCGGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAATGTTTGGATT  
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT  
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT  
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGC  
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT  
GTTTGGGTCAAACAGGTGNTCGCATTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT  
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTAT  
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

09989734 440004  
10004 40004



**FIGURE 61**

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC  
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT  
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTATAGCAACCATAGCC  
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG  
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN  
TGATTGCATTTATGTGGATTNTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC  
CCTGT

**FIGURE 62**

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGCGCGG  
CGNGACACCGGGTTCCGGGAACCATTCACGACGGGGTGGACTGACCTGAAAAAATGTTTG  
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT  
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT  
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA  
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA  
GGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG  
ATNTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG  
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG



**FIGURE 64**

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN  
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG  
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ  
EEAIIIDKYIKLVKVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ  
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA  
GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP  
GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH  
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAEKKGKRGSDYETFKNSFVEA  
SMSVVLKLFPPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI  
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN







**FIGURE 68**

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAG  
SAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



CCGCCCCCGCCCGAGACACCGGCGCCGGGGCGCGGGGGCGCGGGGATGCGGCGCCCGGGGGCGG  
CGATGACCGCGGAGCGACAGCCGCGGGGCCCGGCCCTGACCCCGCCGCCCGCCCGCTGAGCCC  
CCCGCCGAGGTCCGGACAGGCCGAGATGACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC  
CGCTGCTGCTGGGGGCCTTCCCACCGGCCGCCGCCGCCGAGGCCCCCAAGATGGCGGAC  
AAGGTGGTCCCACGGCAGGTGGCCCCGGCTGGCGCGCACTGTGCGGCTGCAGTGCCAGTGGA  
GGGGACCCCGCCCGCTGACCATGTGGACAAGATGGCCGCACCATCCACAGCGGCTGGA  
GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC  
GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT  
GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG  
AGGACCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGCGCG  
CGGGTGATCGCACGGCCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCCTGGCCAGCGGGCACCC  
TCGGCCCGACATCAGTGGATGAAGGACGACGAGGCCTTGACGCGCCGAGAGGCCGCTGAGC  
CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGAGGACAGCGGCAAATAC  
ACCTGCCGCGTGTGCAACCGCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA  
GCGGACCCGTTCCAAGCCCGTGCTCACAGGCGACGACCCCGTGAACACGACGCTGGACTTCG  
GGGGACCCACGTCTCTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG  
AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA  
GTTTGTGGTGCTGCCACGGGTGACGTGTGGTTCGCGGCCCGACGGCTCCTACCTCAATAAGC  
TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACACC  
ATGGGCTACAGCTTCCGCGAGCGCCTTCTCACCCTGCTGCCAGACCCAAAACCGCCAGGGCC  
ACCTGTGGCCTCCTCGTCTCTGGCCACTAGCCTGCCCTGGCCCGTGGTATCGGCATCCAG  
CCGGCGCTGTCTTATCCTTGGGACCCCTGCTCTGTGGCTTGGCAGGCCAGAGAAGACCG  
TGCACCCCGCGCCTGCCCTCCCCCTGCCCTGGGCACCGCCCGCCGGGGACGGCCCGCGACCG  
CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGT  
GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCCAGGCCCAGTTGCTGGC  
CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACTCTCACAC  
ACACTCACACGTGGAGGGCAAGGTCCACCAGCATCCACTATCAGTGCTAGACGGCACCGT  
ATCTGCAGTGGGCACGGGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT  
CGAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCCGAGGCAGTCTGTGTG  
TGAGGCATAGCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC  
ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACAGATGCACAGATATG  
CCGCTTGGGCACACAGATAAGCTGCCAAATGCACGCACACGCACAGACAGATATGCCAACA  
TACAAGGACATGCTGCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCTGGACA  
CACACACACACACGGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACA  
CACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG  
ATATTGCCTGGACACACACACACACACACGCGTGCACAGATATGCTGTCTGGACACGCACAC  
ACATGCAGATATGCTGCCTGGACACACACTCCAGACACACGTGCACAGCGCAGATATGCT  
GCCTGGACACACGAGATATGCTGTCTAGTACACACACACACGCGAGACATGCTGTCCGGACAC  
ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCGACGAGATATGCTGCCTGGAC  
ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA  
TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG  
CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT  
GCAGATATGCTGCCTGGACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGAT  
TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA  
TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA  
CACACATGCACACAGGTGCAGATATGCTGCCTGGACACACGAGACTGACGTGAGCTTTTGG  
GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGACTTTT  
CCTTGCTCCACCGTCACTCCCCAACTGTGCCCGCCTCTGTCCCCGCTCAGTCCCCGCCTC  
CATCCCCGCTCTGTCCCCCTGGCCTTGGCGGCTATTTTTTGGCACCTGCCTTGGGTGCCCAGG  
AGTCCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG  
CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCCTTGGA  
TTTATATTTAAGAAATGAAGATAATATTAATATGTAAGGAAGACAGTGGGTGCAGGGAC  
TGTGGTCTCTCTGGGGCCCCGGGACCCGCTGGTCTTTTACGCCATGCTGATGACCAACCCC  
GTCCAGGCGACAGACACACCCCCCACTGTGCTGGTGGCCCCAGATCTCTGTAATTTTA  
TGTAAGTTTGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

MTPSPLLLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM  
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVLDDISPGK  
ESLGPDSSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK  
DDQALTRPEAAEPRKKKWTLSTLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL  
TGTHPVNTTVDFGGTTSFQCKVRS DVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLP TGD  
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLT VLPDPKPPGPPVASSSSA  
TSLPWPVVGIPAGAVFILGTL LLWL CQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS  
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV  
HQHIHYQC

**FIGURE 71**

CCACAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAAACTTCCCAGGGGACCGCATTCCAGAGTCA  
AGTGACTCTGTGAAGCACCACATCTACCTTCTTGCCACGTTCCCACGGGCTTGGGGGAAAGAT**ATGGT**GGGGGACCA  
AGGCCCTGGGTGTTCTCCCTCCTGGTCTCGGAAGTACATCTGTGTGGGGAGACAGACAGTGTCTACCCAGTCA  
GTAAGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGG  
TGAGTGGACAAACATGGTTCAACATCGACTACCCAGGCGGGGAAGGGCGCATATGAGCGGCTGGACGCCATTGCT  
TCTACTATGGGGACCGTGATGTGCCGTCCCTCGCGCTAGAGGCTCGGACCATTGACTGGACACCTGCGGGC  
AGCACTGGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGGAGCAGCGGCCTGGCCA  
GAACTGCTCTAATTACACCGTAGCGTTTCTCTGCCACACAGGATCCCTGCGCCGACAGACAGAGCGCATCTGGA  
GCCACTGGTCTCCCTGGAGCAAGTGCTCAGCTGCCGTGTGGTCACTGGGGTCCAGACTCGCACACGCATTGTGC  
TTGGCAGAGATGGTGTGCTGTGCAGTGAAGGCCAGCGAAGAGGGTCACTGACTGCATGGGCCAGGACTGTACAGC  
CTGTGACCTGACCTGCCAATGGGCCAGGTGAATGCTGACTGTGATGCTGCATGCTGCCAGGACTTCATGCTTTC  
ATGGGGCTGTCTCCCTTCCCGAGGTGCCCCAGCCTCAGGGCTGCTATCTACTCTCCTGACCAAGACGCCGAAG  
CTGCTGACCCAGACAGACAGTGTATGGGAGATTCCGAATCCCTGGCTTGTGCCCTGATGGCAAAAGCATCCTGAA  
GATCACAAAGGTCAAGTTTGGCCCCATTGTACTACAATGCCAAAGACTAGCCTGAAGGCGGCCACCATTCAAGG  
CAGAGTTTGTGAGGGCAGAGACTCCATACATGGTGATGAACCTGAGACAAAGACCGAGAGCTGGGCAGAGC  
GTGTCCTGTGTCTGTAAGGCCACAGGGAAGCCAGGCCAGACAAAGTATTTTTGGTATCATAATGACACATTGCT  
GGATCCTTCCCTCTACAAGCATGAGAGCAAGCTGGTGCTGAGGAAACTCGACGACGACCAGGCTGGGGAGTACT  
TTTTGCAAGGCCCAGAGTGTGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTCAAGCATCTGATGAG  
ACTCCTTTGCAACCCAGTTTCTGAGAGCTACTTTATCCGGCTGCCCATGATTGCTTTTCAAGATGCCACCAACTC  
CTTCTACTGTGACGTGGGACGCTGCCCTGTTAAGACTTGTGAGGGCAGCAGGATAATGGGATCAGGTGCCGTG  
ATGCTGTGCAGAAGTCTGTGGCATCTCCAAGACAGAGAAGAGGAGATCCAGTGCAGTGGCTACACGCTACCCC  
ACCAAGGTGGCCAAAGGAGTGCAGCTGCCAGCGGTGTACGGAAACTCGGAGCATCGTGCGGGGGCCGTGTGCTGCT  
TGCTGACAATTGGGAGGCCCATGCGCTTTGGCCATGTGTACATGGGGAACAGCCGCTGTAAAGCATGACTGGCTACA  
AGGGCACTTTTACCCTCCATGTCCCCAGGACACTGAGAGGCTGGTGCTACATTTTGTGACAGGCTGCAGAAG  
TTTTGTCAACACCACCAAGTGTACTCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAAATCAAGATGCTTCG  
TCGGAAAGAGCCCATCACTTTGGAAGCCATGGAGACCAACATCCCTCTGGGGGAAGTGTTGGTGAAGACC  
CCATGGCTGAATGGAGATTCCATCCAGGAGTTTCTACAGGCAGAATGGGAGCCCTACATAGGAAAAATGAAG  
GCCAGTGTGACCTTCCCTGGATCCCCGGAATATTTCCACAGCCACAGCTGCCAGACTGACCTGAACTTCATCAA  
TGACGAAGGAGACATGTTCCCCCTTCCGAGCATATGGCATGTCTTCTGTGGAATTCCAGATGAGGTCACTCAG  
AGCCACTTAATGCTGGCAAAGTGAAGTCCACTTGACTCGCCACAGTCAAGATGCCAGACACATATCCACA  
GTGAAACTCTGGTCACTCAATCCAGACACAGGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAAATCAAAG  
GAGGAACAAAAGAGAAGACAGAACCTTCTGTGGTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTCTTTAACTGG  
ATGTTCTGTAAGAGAGCGGTTGCTTTGTTAAGTGAAGGCCCTACCGAGTGAGAGGTTCTTGCTTACTGAGCAG  
ATCCAGGGGGTTGTGATCTCCGTGATTAACCTGGAGCCTAGAACTGGCTTCTTGTCCAACCCTAGGGCCTGGGG  
CCGCTTTGACAGTGTATCATCAGAGGCCCAACGGGGCCTGTGTGCTGCTCTGTGTATGACCACTTCCCCGTATG  
CCTACTCTGCCTATGTCTTTGGCAAGCCTGGCTGGGAGGAAGTCAAGACAGTGGAGTCTTCTGCTAAATTC AAC  
CCAAATGCAATTGGCGTCCCTCAGCCCTATCTCAACAAGCTCAACTACCGTTCGGACGGACCTAGAGGATCCACG  
GGTTAAAAAGACAGCTTTCCAGATTAGCATGGGCAAGCAAGGCCCAACTCAGCTGAGGAGAGCAATGGGCCCCA  
TCTATGCTCTTTGAGAACCTCCGGGCATGTGAAGAGGCACCAACCGATGCAGCCCACTTCCGGTTCTACAGATT  
GAGGGGGATCGATATGACTACAACACAGTCCCCTTCAACGAAGATGACCTATAGCTGGACTGAAGACTATCT  
GGCATGGTGGCCCAAAGCCGATGGAATTCAAGGCTGCTATATCAAGTGAAGATTGTGGGGCCACTGGAAGTGA  
ATGTGCGATCCCGCAACATGGGGGCACCTATCGCGGCAGCAGTGGGGAAGCTGTATGGAATCCGAGATGTGAGG  
AGCACTCGGGACAGGGACCAGCCCAATGTCTCAGCTGCCTGTCTGGAGTTCAAGTGCAGTGGGATGCTCTATGA  
TCAGGACCGTGTGGACCCGACCTGGTGAAGGTCACTCCCCAGGCGAGCTGCCGTGAGCCAGTGTGAACCCCA  
TGCTGCATGAGTACTGCTGCCACTTGCCACTTGCACTCAACAACGACACCCAGTGAAGTACACCATGCTGGCA  
CCCTTGGACCCACTGGGCCACAACATATGGCATCTACACTGTCACTGACCAGGACCTCGCACGGCCAAAGGAGAT  
CGCGCTCGGCGGGTGCTTTGATGGCACAATCCGATGGCTCCTCCAGAAATCATGAAGAGCAATGTGGGATAGCCC  
TCACTTTCAACTGTGTAGAGAGGCAAGTAGGCGCGCAGAGTGCCTTCCAGTACCTCAAAGCACCCGAGCCAG  
TCCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCCTCGAGGAGGACGAGCAGCGAGCAGGAGTGGCCAGGCT  
CCAGGGTGGAGTGGTGGCTCTCTGAGATTTCATAGAGTTGCTCAACAGCCCCCTGATCAACT**TAAGT**TTTGTGGT  
ACTTCAACCCTCTTCTGCCCTCATTTCACTGTGACAGCCATTGTGAGACTGATGCACAAACTGTCACTTGGTTAAT  
TTAAGCACTTCTGTTTTCGTGAATTTGCTTGTGTTTCTTTCATGCCTTTACTTACTTTGTGCCCTGCTACTGA  
TTGGCAGCTGGCCCCACAATGGCACAATAAAGGCCCTTTGTGAAACTGTTCTTTAAATGAACACAAGAAAT  
GGCCACTGGTAAAACTCTGCAGCTTCAACTGTACTTCATTTAATGCCATTAAATGCAAATATACTTCTCTCTCT  
TTTTGCATGGTTTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAAATACCAATATAAAGCATAT  
TTCTTGGCCTTGCTCCACAGGACATAGGCAAGCCCTTGATCATAGTTTCATACATATAAATGGTGGTGAATAAAG  
AAATAAAACACAATACTTTTACTTGAATGTAAATAAATACTTATTTATTTCTTTGCTAAATTTGGAATTTCTAGTGC  
ACATTCAAAGTTAAGCTATTAATAATAGGGTGATCATAGTTCCTCTACCAAGTCTGGAAAGAACATCTCCTGGT  
ATCCACAATACACAGGTTGCTAECTGATTTGTACATTTCCCTTTGCACTTCGCTTTTGTCTTGTGCTAGAAAC  
CCAGTGTAGCCAGGGCAGATGTCAATAAATGCATACTCTGATTTTCGAAAAA

**FIGURE 72**

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI  
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ  
RPGQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS  
EASEEGQHCMGQDCTACDLTCPMGQVNADCDAQDFMLHGAVSLPGGAPASGAAYLLTK  
TPKLLTQTDSDGRFRIPLGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY  
MVMNPETKARRAGQSVSLCCKATGKPRPDYFWYHNDTLLDPSLYKHESKLVLRLKQQHQAG  
EYFCKAQSDAGAVKSKVAQLIVTASDETFCNPVPESYLIRLPHDCFQATNSFYDVGRCPV  
KTCAGQQDNGIRCRDAVQNCCKGSKTEEREIQCSGYTLPTKVAKECSCQRCTETRIVRGRV  
SAADNGEPMRFGHVYMGNSRVSMGTGYKGTFTLHVPQDTERLVLTFFVDRLQKFVNNTTKVLPFN  
KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV  
KASVTFLDPRNISTATAAQTDLNFINDGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL  
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL  
DVPESRRCFVKVRAYRSEFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA  
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNRYRRTDHEDPR  
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN  
EDDPMSWTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS  
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTLKVIPOGSCRRASVNPMLHEYLVNHLPLAV  
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSBGSSRIMKSNVGVALT  
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGRQGGVVASLRFPRVA  
QQPLIN

**FIGURE 73**

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCTCAATATACCTGAATACGCAC  
AATATCTTAACCTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG  
ACCTACCTACCCGTACGCATACATACATATGTGTATATATATGTAACTAGACAAAGATCGCAGATCATAAAGC  
AAGCTCTGCTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAGATGATTTGTCAAGATCCCTGTCGATTTCATG  
CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT  
CAGATTTACACGGAAGAAGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA  
TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC  
CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCTGAGCTGATGTTTGAATTTT  
GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT  
CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC  
AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA  
GATGCTTTTTACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA  
AGAGTACTCAACAGGGTATACAACAAAATAGCAAAAATAATCCACTTTGAAATCAAAGACAGGTTCCGCGTTTTTG  
CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATAACAACCAAGAAACTCAGAGATTTCTTT  
ACAGTACAGACCTGAGGATAAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC  
ACGCTACTTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAACTCTCCATGCCACTGTATGTG  
TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG  
AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC  
CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA  
ACAACGTGCGCTGCCTGTGCCCCGCCGATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC  
AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCCGCACGGCACCCCAAGCGCTGCTGCTGCTGACCACGCTGCT  
GGGAACCGCCAGCCCCCTGGTGTCTAGGTGTACCTCCAGCCACACCGGACGGGCTGTGCCGTGGGGAAGCA  
GACACAACCCAAACATTTGCTACTAACATAGGAAACACACATACAGACACCCCACTCAGACAGTGTACAAA  
CTAAGAAGGCCTAACTGAACCTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTT  
TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA  
TTGGAAAGGCTGCGACAGCCCCCAACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTGGC  
TACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCCAAGTGTGTGGACCAACCAATAGCATTCTTTGCTGTGAG  
GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC  
CTGTGCTTTAGTGAACGTTGCTCTGTAACCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACA  
TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA  
GCACACACCCACTATACAAGAGTGGCTATAGGAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT  
ATTTTCTTGAACCTACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT  
GTATCTAATTCGAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTCGTTCTGTTTTGTTTCACTGTGCAGA  
GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGTAAATAAGA  
TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA  
TAAACGATTTTCAAGAAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACT  
TTACTGATTTCTGTGTGGACTGAGTACATTGAGCTGACGAATTTAGTTCCCAGGAAGATGGATTGATGTTCACT  
AGCTTGACAACTTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAAA  
AAAAAA

**FIGURE 74**

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK  
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK  
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMLEKSLDYGRTWQPYQYYATDCLDAF  
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD  
TTKKLRDFFTVDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN  
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH  
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGGAPPHGTPALLLLTTLLGTAS  
PLVF

**FIGURE 75**

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGATTGCTGAGGAGGCGG  
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGGTCCGGGCGAGGTGTCCTCATGACTT  
CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAAGGGATGGACTGCC  
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGGAAATGGAGGAGACGGCTCA  
AGAGTTTAGCCTTGCGACTGGCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT  
AGTATACATTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC  
AGCAGCCATGGCCTTCTGCTTCTTGAGACCCTGTGGTGGGAATTCACAGCTTCTATGACA  
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTTCTTGAGTTTGACAGCATCATTAG  
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAAAAAAT  
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA  
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA  
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT  
CATTCGAGGAGTTCACCTTGCGAAGCATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT  
TGGACCAAACCTCG**TGA**GCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT  
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA  
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAAACATCTGCCCATGTGTATTG  
ATGGCAGAGCTGTTGCCCAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT  
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT  
GCTTCAGAAGTGTTATTTTATGAATCATTATGATTTGATCCCCCAGGATTCTATTTTGT  
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT  
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTATGTTACACAAGCTTCTTACGGTTTTT  
TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC  
TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAATG  
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAATAAAAAAAAAAAAAAAAAA  
AAAAG

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF  
SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQVKW  
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA  
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS



FIGURE 77

[illegible]

**FIGURE 78**

CTCAGCGGCGCTTCTCTCGTAGCGAGCCTAGTGGCGGGTGTTTGTCATTGAAACGCTGAGCGCGA  
CCCCGACCTTAAAGAGATGGGGAGCAAAGGGAGGACAGAGACCCTTTAAACGAGGCGGGTGCTG  
CCTGCCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT  
TCTGTGCGCAGGCTGCGAGGAAAGGCCCTTAGGCTGGGTCTGGGTGCTTGGCGGCGGGCGCTT  
CCTCCCCGCTCGTCTCCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA  
**TGGA**AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC  
GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC  
CCGCTTCAAGAAGCCTGCTGAGTTACACACAGTGGATGATGAAGATGCCACCGTCAACAAGA  
TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTCTCCTGCC  
TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAATACTACTACAGTGGCT  
CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTCCCCAACCTGTCCCTCA  
TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG  
GGTGTCTCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT  
AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT  
ATGACTTTTGGGAGTACTATCTCCCCTACCTCTACTCATGCATCTCCTTCTGGGGTTCTG  
CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT  
AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG  
CAGCCCTGACCCGCAGGATCTGTAATCCTACTTCTGCTGGCTGCCTTTAGACATGGAGCTG  
CTACACAGACAGGTCTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGG  
TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGCTGGTGCTGACGG  
GCCTGTCTGTGCTCATTTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG  
CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG  
TGCCGTCAATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTATA  
GCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACACTGCCATGACGCAGATAATT  
GGGAACATGTGTCTGTCTCCTGGTCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCTGGG  
GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA  
TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC  
ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT  
CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCA**GTGA**CCTCCAGCTGGGGGTGGGA  
AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG  
ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT  
GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT  
CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCTGATCCCAAATCTGTTTACACATCA  
ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT  
AGGGTGGGGTAGGGGCAGGGAAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC  
CTTGCTCTGGCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG  
AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA  
AAAAAA

**FIGURE 79**

MEAPDYEVLSVREQLFHERIRECIIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK  
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL  
IFLMPFAYFFTESEGFAGSRKGVLGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL  
YDFWEYYLPYLYSCISFLGVLLLLLVCTPLGLARMFSVTGKLLVKPRILLEDLEEQLYCSAFEE  
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLKRRKASAWQRNLGYPLAMLCLLVLT  
GLSVLIVAIHILELLIDEAAMPGRMQGTSLGQVSFSKLGSFGAVIQVVLI FYLMVSSVVG FY  
SSPLFRSLRPRWHD TAMTQIIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRFNWLG NFY  
IVFLYNAAFAGLTTLC LVKTF TA AVRAELIRAFGLDRLPLPVSGFPQASRKTQH Q

**FIGURE 80**

GGCTGCCGAGGGAAGGCCCTTGGGTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC  
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC  
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA  
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC  
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 81**

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTGC  
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC  
TGTCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT  
CCCCGTTGTCNTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA  
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT  
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC  
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC  
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCCTTCT  
CCATCATCAGCAATGAGGTGCTGCACTCCC

**FIGURE 82**

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT  
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTT**CATG**CTGCTGT  
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT  
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA  
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA  
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTTCAGGAATCTGGAGAGTACAGATGCCAG  
CCCCAGGGCTCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC  
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCC**TAGG**CCCTCTC  
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC  
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA  
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA  
AAA

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL  
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDIFSSEMGFPHAAQANVELLGSSDLLT

Table 1. (continued)	
10	100
11	110
12	120
13	130
14	140
15	150
16	160
17	170
18	180
19	190
20	200
21	210
22	220
23	230
24	240
25	250
26	260
27	270
28	280
29	290
30	300
31	310
32	320
33	330
34	340
35	350
36	360
37	370
38	380
39	390
40	400
41	410
42	420
43	430
44	440
45	450
46	460
47	470
48	480
49	490
50	500
51	510
52	520
53	530
54	540
55	550
56	560
57	570
58	580
59	590
60	600
61	610
62	620
63	630
64	640
65	650
66	660
67	670
68	680
69	690
70	700
71	710
72	720
73	730
74	740
75	750
76	760
77	770
78	780
79	790
80	800
81	810
82	820
83	830
84	840
85	850
86	860
87	870
88	880
89	890
90	900
91	910
92	920
93	930
94	940
95	950
96	960
97	970
98	980
99	990
100	1000

[illegible][illegible]



**FIGURE 85**

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ  
 AQEKFQDLGAAYEVLSDSEKRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQQ  
 DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ  
 MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFIGEGEPHVDGEPGDLRFRIKVVKH  
 PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD  
 NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

**Important features:****Signal peptide:**

amino acids 1-22

**Cell attachment sequence.**

amino acids 254-257

**Nt-dnaJ domain signature.**

amino acids 67-87

**Homologous region to Nt-dnaJ domain proteins.**

amino acids 26-58

**N-glycosylation site.**

amino acids 5-9, 261-265

**Tyrosine kinase phosphorylation site.**

amino acids 253-260

**N-myristoylation site.**

amino acids 18-24, 31-37, 93-99, 215-221

**Amidation site.**

amino acids 164-168

**FIGURE 86**

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA  
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT  
CCGCAGAACNTGAGCACNTTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG  
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTA AAAAGG  
CCTATAGGAACTAGCCCTGCAGNTTATCCCGACCGGAACCCTGATGATCCACAAGCCAG  
GAGAAATTCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA  
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT  
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA  
AATATTCCAAGAG

GGCACGAGGCGGCGGGGAGTGCGCCGGGAGCCACAGCCTGAGGGCCCTCAGGT  
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA  
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**  
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC  
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC  
TAAGCCCATTTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC  
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC  
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC  
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA  
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG  
TACCCTCCGTTGGACCCCAAACCTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCAG  
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC  
AGTCTCTGTGCGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG  
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAGT**GCCT  
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT  
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG  
TAAAGCAGGAGATCCCCGTCAAGTTTATGCCTCTTTTGCAGTTGCAAACCTGTGGCTGGTGAGT  
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA  
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT  
AAGAAATCAAGAGGTTTACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTGAGAATG  
TGTGGCAATTCTGATCTGCATTTTCAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTTC  
TTCTTTTGGCAAGACTTGTACTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT  
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAGGTTTGGGTTTGAAGCTGAGGAAC  
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG  
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAAAAAA

METVIVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSESEL  
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS  
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI  
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

**FIGURE 89**

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC  
CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA  
TTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT  
TGGAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT  
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC  
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTCAG  
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT  
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA  
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA  
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG  
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA  
ACTAAGAAGAAGTCAGCAAGCAAACCTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA  
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA  
CTGTGGTGCCTGTTTCTTTTCTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT  
TTTTAGAAGTGTCCTACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA  
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG  
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

**FIGURE 90**

MISLTDQTQKIGMGLTGFGVFLLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK  
HKMKATGFFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIIRVPVLGSLLNLPGI  
RSFVDKVGESNNMV

**Important features:**

**Transmembrane domains:**

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

**N-myristoylation sites.**

amino acids 11-16, 51-56 and 116-121

**Aminoacyl-transfer RNA synthetases class-II protein.**

amino acids 49-59

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC  
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA  
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTCTGTCTTTGGA  
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT  
GGCTTTTGTAATTGTTTTAGAAAGAACATTTCAGATTCTTCTTCCAAAACATAAAATGAAAG  
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG  
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

**FIGURE 92**

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA  
GGCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA  
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC  
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC  
ATCAACACCATTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA  
GATCAACTGCAGACTGTCTTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT  
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT  
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA  
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA  
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT  
CGCAAGACGGTTGCCACCAAGTTTGACGACCTCCGGGACTACCCCGAGAAGTATTTTTTCT  
GATTCAGTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC  
GGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC  
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTAGAAA  
TAATGAAAATCCAACACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATG  
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC  
AAGCTCTACCAGGAGAAGGATGCCTTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA  
GACGCCCATGGTGCCCCCGGCGGCCCTGGACCCTCGTGAAGTGGCTGTTTTGGGCCTCGC  
TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG  
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT  
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGAC**T**  
**G**ACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAAGTGGTGGCCTCTGCATATCCT  
CCTTAGTGAGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA  
CCTCTCCAGCCAGGAGTCTGGTCTCAAGGCCGATGGGGAGGAAGATGTTTTGTAATCTTT  
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGC  
TGTGTGGTGAGTGTGAACCTTTGTTCTGTGATCATAGAAAGGTATTTTAGGCTGCAGGGGAG  
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT  
AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC  
CAAGAAAAAAAATTAAGTGCTTTTCTGGGTCAAAAAAAAAA



**FIGURE 93**

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV  
MLEWWSGTECTIFTDPRAYLKYGKENAIVVLN HKFEIDFLCGWSLSERFGLLGSKVLAKK  
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQH LRDYPEKYFFLIHCEGTRFTEKKHE  
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRN VVS AVYDCTLNFRNNENPTLLGVLNGKK  
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN  
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS  
KQKLND

93/330

**FIGURE 94**

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG  
GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA  
GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA  
TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG  
GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT  
TGGTACAAATTCCGTCGTCATTGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA  
AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA  
TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT  
TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACCTGGGTATATAAAC  
TGTATCAGGTTCCGTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT  
TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA  
CAAGAGGAATTAAAGAGTATATGCAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT  
AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG  
CAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA  
CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTTCATGTGTTATGTCTTTAAAAAATAGACA  
TGTTTCTAAAAGTAGCTGTAACCTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA  
TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAGCAT  
AAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTAGATACACAAGA  
CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAATGAGCAGCC  
CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA  
TTTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT  
TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC  
CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC  
ATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACCTAAGTTGTTGAGGGGAAGGCT  
TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA  
TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC  
TGGGCAACGTATTGAGACCATGTCTATTAATAAATAAATGGAAGCAAGAATAGCCTTAT  
TTTCAAATATGGAAGAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAG  
TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCA  
ATAAATTTGCAAAACATCATCTAAAATTTAAAAAAAAAAAAAAAAAAAAA

MEGESTSAVLSGFVLGALAFQHLNTDS DTEGFLLG EVKGEAKNSITDSQMDDVEVYTI DIQ  
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH  
FSNQDLVFLLLTPSIITESCSTHRLEHS LYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC  
MSTGFSRAVQTHSSKFF EEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN  
RLKREIEKRGAQIQAA REKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHSVSKSS  
CNYNHHLDVDNL TLMVEHTDIPEASPASTPQIIKHKALDLD DRWQFKRSRLLD TQDKRSKA  
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

**FIGURE 96**

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC  
 CCAAGCAGCGCGCAGCGAACGCCCCGCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCT  
 CCTTCCCCGCGTCCCCGCTCGCCGCGCAGTCAGCTTGCCGGGTTGCTGCCCGCGAAACCCCGAGGTCACCA  
 GCCCCGCGCCTCTGCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCCTGGCCCGCGCCTGGCACC  
 GGGGACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTCCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCAC  
 CAACTCCAACCTCCTTCTCCCTCCAGCTCCACTCGTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAG  
 CGCCGCTTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGCGCCCGCACC**ATG**GCACGGTTCGGCTTGCC  
 CGCGCTTCTCTGCACCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGG  
 AAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCAT  
 TTGAAGATCTGTCCCCAGGGTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA  
 TGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTGCTTCACGTTACAAGAAGTTTG  
 ATGAATTTCTCAAAGAACTACTTGAAGATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT  
 TTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT  
 GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACCTCCAGT  
 ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT  
 CGCAAATTTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGG  
 AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGATCCCATGCCCTGTTGAAGATGATCTACT  
 GCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGCTGTTTG  
 GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA  
 GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG  
 ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATT  
 TCTCGTTCCATCTCTGAAAGTGCTTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACACAGC  
 AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCCT  
 CCCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG  
 AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA  
 GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA  
 AGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGT  
 GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA  
 GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTCTCTGG  
 TTATGCAGAGAGAGTGGAGAT**TAA**TTCTCAAACCTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAAGTT  
 ATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC  
 CACTGGTTTAAAGAAGTGCTGACTTTGTTTTCTCATTAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT  
 TCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA  
 TTTTATCACTCTATTATTTGTTTGTATGTTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAACTGTGATCT  
 CGCCTTGTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA  
 GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAGAAAAAGCCCAAAAGC

**FIGURE 97**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ  
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF  
VKTYGHLYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFLVNSQYHFTDEY  
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL  
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIES  
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGCQPPKPLPAGRISRSISESAFSARFRPHHPEE  
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYL  
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE  
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**FIGURE 98**

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT  
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA  
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGC  
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC  
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA  
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC  
AGTGCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA  
AAGCCAAACAAGCATTCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT  
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA  
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC  
CCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC  
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT  
AATTACCTGAAAGATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTTAACCTTAAATGC  
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

**FIGURE 99**

MKVLISL L L L L L P L M L M S M V S S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R  
K F M T V S G L P K K Q C P C D H F K G N V K K T R H Q R H H R K P N K H S R A C Q Q F L K Q C Q L R S F A L P L

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 100**

A**ATG**GCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT  
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA  
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA  
GTTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA  
AACATTCATCAAAG**TGA**CATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA  
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC  
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC  
CTGACTGCATTTTTGCTTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG  
ATGGAGAGGAAA





**FIGURE 102**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT  
CAGAGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC  
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCC  
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG  
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG  
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC  
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT  
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT  
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCC  
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA  
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGAAACACATTGGGGATGGCTG  
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG  
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC  
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA  
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTAGCAGAATGAGAGAAGACATT  
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC  
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC  
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC  
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT  
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT  
CTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT  
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA  
AACCACG

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELF SQI  
KGLTGASGKVALLELGC GTGANFQFYPPGCRVTC LDPNPHFEKFLTKSMAENRHLQYERFVV  
APGEDMRQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRVLRPGGVLF FWEHVAEPTYGSWAFM  
WQQVFEPTWKHIGDCCLTRETWKDLENAQFSEIQMERQPPPLKWL PVGPHIMGKAVKQSF P  
SSKALICSFPSLQLEQATHQPIY LPLRGT

**FIGURE 104**

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAAGTTGCCTCATCGCAGG  
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA  
ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTGG  
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG  
TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA  
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCCTGTGGTCATCGCTGCATCTG  
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTTCAGCACAACACTCGCTCCAAT  
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG  
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA  
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC  
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA  
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG  
AAGATTGTGATTTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC  
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG  
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA  
ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT  
GGTAGCATCACAAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC  
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA  
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT  
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTC AACCTAATCCGAAG  
ATATACCGAGATCTCAAACATAAAGTGAAACAGAATTTGAACTGTAAGCAAGCATTTCTCAG  
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA  
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC  
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT  
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA  
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT  
ACATTTTTTC

**FIGURE 105**

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTD SGIVGPQPIDFVPNALRHAVDGR  
QEEIPVVIAASEDRLGGAIAAINS IQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK  
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT  
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA  
NLTEWKRQ NITNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS  
SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWI PDPTGKFNLIRRYTEISNIK

105/330

**FIGURE 106**

TGGTTTTTGCCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT  
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT  
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT  
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC  
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA  
TACAAAATTGTCAATTTTGACCCTAACTTTTGGGAAGGAAAAGTAAAGGAGGATCCTGACCA  
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTCTACTTGCCAATTCTGGTTCCCAGCG  
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC  
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC  
TAAAGTTGTCATCCGTGGAGCAGGAAA

**FIGURE 107**

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG  
 GAGCTAGAGGGCAAGTGTGCTCGGCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC  
 TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA  
 TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT  
 TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT  
 TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT  
 CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC  
 CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTGAGGAATTGTAG  
 GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA  
 GGAGATTCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA  
 ACAGCATTGAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA  
 GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG  
 TCAATTTTGACCCTAAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC  
 ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG  
 CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA  
 CTGAAGCCAGGACATGCAGCTGCATTTTTCAGAAGATTGTGATTGAGCCTCTACTAAAGTTGT  
 CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA  
 TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC  
 CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT  
 AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG  
 TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT  
 GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA  
 TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA  
 GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA  
 CAGAATTTGAACTGTAAGCAAGCATTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG  
 TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG  
 GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT  
 ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA  
 CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT  
 GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAA

**FIGURE 108**

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT  
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG  
TCTGTGCTGGTCTGAGGGTGCTGCCTGTCAATGGGGGCAGCCATCTCCCAGGGGGCCCTCATC  
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCACTCCTCTGCTG  
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC  
TGGCCCCCTGTCCTGAGAAGGCCCCACCACCCAGAACCCAGCCATGAAGGCAGCTACCTGC  
TGCAGCCCTGAAGGCCCTGGCCTAGCCTGGAGCCCAGGACCTAAGTCCACCTCACCTAGAG  
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT  
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC  
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT  
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTTCCCTAGGCTGAGCACC  
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG  
GCTGGGCCCCCTCCCCTGGTCCTCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA  
AAAAAAAAAAAAAAAAAAAA



109/330

**FIGURE 109**

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRPHH  
PRSPAMKAATCCSPEGPWPSLEPRT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 110**

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA  
GTTCCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA  
CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAA  
TC**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC  
ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTAT  
TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC  
TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTA  
TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT  
GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC  
AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTG  
AGCCTGGGAAGTGCAGGAGCTGCCAGGTTATGGAAGGCGCCAAGTGGAATATAAGCCCCCT  
TTCGGGCATTTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA  
TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTACTTGTTATTTCAACAGAAGT  
AAAAATGATCCTCCTGATCATCCCATCCTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCA  
AGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAGGATTCGAGAATCATTTGTCA  
TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCAATTGTCCAGGTACCTGTTCCGA  
TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA  
TACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT  
TGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTTCTA  
GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG  
GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCC  
ATAGTTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT  
CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT  
CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA  
ATGAGGAGGGAACAGAAGTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCT  
GGAAAACATTTCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT  
AGTGAATTTTTTTTTTAAAGACCTAATAAACCTATTCTTCCTCAAAA

Downloaded from www.oxfordjournals.org/ by guest on July 10, 2015

**FIGURE 111**

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYYDYTNDL  
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ  
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI  
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM  
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKIL  
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH  
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN  
EEGTELQAIVR

[illegible]

**FIGURE 113**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC  
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL  
SLPRWRESFIVLESKPKKGVTPSALTYSSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ  
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQMDLWSTATYTSSQNRPRADPGIQ  
RQDPSGAAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR  
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNRDLKTAIEKITQRGGLSNV  
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE  
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLOPLVKRVCDTDRACSKTCLNSADI  
GFVIDGSSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLFEFGFDKYSSKPD  
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKG  
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFNLHQYVPRIIQNICTEFNSQPRN

[illegible][illegible]

**FIGURE 115**

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE  
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL  
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM  
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK  
AIRNYTRFDDWYLVWQMYKGTVSMPVFQSLEAYWPGQLSLIGDIDNAMRTFLNYYTVWKQFG  
GLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG  
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY  
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT  
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG  
GGCAGAAAGGAGGGTGCTTCGGAGCCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA  
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT  
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA  
**AATG**CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT  
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTC  
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA  
AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT  
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC  
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG  
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA  
TCACCAAAGATGGCTTCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC  
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG  
GGGTATTCCAGTGACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA  
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA  
GGAGAGGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT  
GGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG  
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG  
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT  
CTCA**TAG**GTTTGCGBAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC  
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA  
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACAC  
TGA CTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC  
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC  
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA  
TACACCCAGCACTTGCAAGGCTAGAGGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG  
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT  
GGCTTGAGAGCCCCACTTTCCCGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG  
TGTTGAGTTCACCTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT  
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC  
CATGA ACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT  
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTTGTGCTCCTTTTTTTC  
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA  
AAAAAAAAAAAA



## **FIGURE 117**

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLIMWSPVIAPGE  
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDDITATVPYNLRVRATLGSQTSW  
SILKHPFNSTILTRPGMEITKDGHLVIELEDLGPQFEFLVAYWRREPGEAEHVKMVRSG  
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMILV  
VVPLFVWKMGRLQLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

### **Important features:**

#### **Signal peptide:**

amino acids 1-29

#### **Transmembrane domain:**

amino acids 230-255

#### **N-glycosylation sites.**

amino acids 40-43 and 134-137

#### **Tissue factor proteins homology.**

amino acids 92-119

#### **Integrins alpha chain protein homology.**

amino acids 232-262

**FIGURE 118**

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT  
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA  
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG  
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT  
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC  
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA  
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCGAATACCAGGGGGAGTACGAGAGCCT  
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG  
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC  
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC  
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG  
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG  
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC  
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

**FIGURE 119**

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGCGTG  
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG  
TCAACATCCGGGGCAAACCTGGTGTGCTGAGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG  
AATGTGGCCAGCGAGTGC GGCTTACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG  
AGACCTGGGCCCCCACCCTTTAACGTGCTCGCCTTCCCCTGCAACCAAGTTTGGCCAAACAGG  
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGACCTACAGTGTCTCATTCCCC  
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA  
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG  
TGGTAGGGGCTTGGGACCCAACTGTGTGAGTGGAGGAGGTGAGACCCCAAGATCACAGCGCTC  
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**ACCACCGCGTCTCCTCCTCCACCA  
CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG  
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA  
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACCTCCTGGCCAATGAG  
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA  
TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC  
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC  
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT  
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT  
GTTATTTCTCCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA  
AACAATACCTCACGATATAAAATAAAAAATGAAAGTATCCTCCTCAAAAA

120/330

**FIGURE 120**

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLVKYRGSVSLVVNVASECGFTDQ  
HYRALQQQLQORDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG  
AHPAFKYLAQTSGKEPTWNFWKYLVPDQKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CGGACGCGTGGGCGGGCCGGGACGCGAGGGCAAAAGCGAGCC**ATG**CTGTCTACGTCGGGATGCTGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGGCCCGGCCGCCCTCTCTCGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCTCAGTTCCAGAGAGGTGGATCGCATGGTCTCCACGCCCATCGGAGGGCCTCAGCTACGTTCAGGGGTGCACCAAAAAGCATCTTAACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCCTTGGTCTGTCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAAAGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGACCTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCAGGCGGGCATCATTCTGGTGCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCTCAAGAAGGTGGGCTGCAAGGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCTGTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCCTTGAAGAGTCAGAGGCTCCCAGATCTGACCACAGTCATCTCGGTGGATGCCCCCTTTGCCGGGGACCCTGCTCCTGGATGAAGTGGTGGCGCTGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTTCTGTCTGCCATGACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCCTCTCCACTACAACATTGTCAACAACCTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTGATGTACGGTGCCACACCTCATCCTGGCCTCTCCATCTTCAATGGCAAGAAGGCACTGGAGGCCATCAGCAGAGAGAGGAGGCACCTTCTGTATGGTACCCCCACGATGTTCTGTGGACATTCTGAACCCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCAATTGCTGGGTCCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGCATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCAAGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAGCTCGAGGACTTCTTTACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGACGATCGGATGGGGGAAGAGATTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGACCACGGTGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCGGAAGTACATGTGTTTGTACAAACTACCCCCTCACCATTTTCAGGAAAGATCCAGAAATTCAAACTTCGAGAGCAGATGGAACGACATCTAAATCTGT**G**ATAATAAGCAGCAGGCCTGTCTTGCCGGTTGGCTTGACTCTCTCCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTCTGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAACTAAGAGCTCCTGGATGGGTCCGGAACTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCTGCCCAGGCCCTCCCTCCTGTCCATCCCCCACATTCCCCTGTCTGTCCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTTGAAAAAAAAAAAAAAAAA

**FIGURE 122**

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVD RMVSTPIGGLSYVQ  
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEV DKAASGLLSIGLCKG  
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY  
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN  
QQFLSCHDPINIQTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY  
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY  
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG  
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT  
MNEQGFCIKIVGRSKDMIIRGGENIYPAELEDDFFHTHPKVQEVQVVGKDDRMGEEICACIRL  
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

**Signal Peptide:**

amino acids 1-22

**Transmembrane Domains:**

amino acids 140-161, 213-229, 312-334

**Putative AMP-binding Domain Signature:**

amino acids 260-271

**N-myristoylation Sites:**

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,  
314-319, 318-323, 379-384, 380-385, 409-413

**N-glycosylation Site:**

amino acids 282-285

**FIGURE 123**

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA  
TGATCCTGCCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG  
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC  
CATCAGCAGAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTTCGTGGACATTCTGA  
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC  
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT  
TGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGG  
AGCAGAAGGCAGAAAGCGTGCGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG  
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT  
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT  
ATTGGACAGGAGATGTCGCCAC

**FIGURE 124**

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC  
AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG  
TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGG  
GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC  
AGTGCGGGGTTGCGGTTCTGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGG  
TTCTGGCGTTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC  
ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGAGAGTG  
CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCTGTGAGCT  
GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA  
GCTAATGTGACTGTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA  
TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT  
CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT  
CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCAACCACTTCTACCTCGGCCCCAGTGAG  
ACCCACATCCACCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG  
AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC  
CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG  
TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTACTGT  
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT  
CATCACTTCCTGTTCCCACTGGACTGGGCTGGCCAGCCCCTGTTTTTCCAACATTCCC  
CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTGGGAAATAAAATACCGTTGTATAT  
ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC  
TCCGCTTGTCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG  
AGAGAGAGGATGCTAAGCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG  
GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCCTACTCCCCGCATCTTTGGG  
GAATCGGTTCCCATATGTCTTCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC  
CCAATTCGCCCTATAGTGAGTCGTA



125/330

**FIGURE 125**

MDPARKAGAQAAMIWTAGWLLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT  
EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL  
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV  
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT  
VASTTSVTTSTTSAPVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGAAGHQDRSNSG  
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 126**

CGGGACTCGGCGGGTCTCTCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCCATGGAGT  
TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG  
GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT  
CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT  
TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC  
AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA  
CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCTTAAGAAAACCGGCCACTTC  
AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTA  
ACACCATTCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCTT  
GCCACCTCCCGTGATGTGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTG  
GTCTTTGTGGCTACTTGTTTGTGGATGGTATTGTGTTTGTTAGTGAAGTGTGGACTCGCTTT  
CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCAC  
CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC  
GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC  
TTTGTGATCATTCCTTCATGGACTCCTTTCACTCCTTTAACAAAAACCTTGCTTCCTTATCCC  
ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC  
CAGCGTTGACGTGAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG  
AGGAGTCCCCATCTGCCCCGCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT  
ACTCTGCCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG  
GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG  
ATGCAGACTACAGTCCCTGCAATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCC  
GTTGGGGCCAGCACACCGGGATGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG  
TCCCCTTAGATGGGCAGCAGAGGCAACTCCCGCATCCTTTGCTCTGCCTGTGCGTGGTCAGA  
GCGGTGAGCGAGGTGGGTTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG  
GTTGAAGGTCATAACGAGAGTGGGAACCAACCCAGATCCCGCCCCTCCTGTCTCTGTGTT  
CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT  
CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

127/330

**FIGURE 127**

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK  
CSFNQKPRAPGDEEAQVENLITANATEPQKQRTevQPSGGSLWNLRRLLEPLDANVDA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 128**

AAACTTGACGCCATGAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT  
CCTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG  
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT  
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT  
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT  
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT  
CTTTCCTGCCTCAGGAAGCTCCAATAAAACATTTTCCATCCAAA



**FIGURE 130**

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCCATGGGGGTGGAGATTGC  
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC  
TTCTCCAGCCAGTTCCAATCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC  
TTTGCCGGCCACTCATGAGAGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCT  
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA  
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT  
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTATTATC  
TTCATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT  
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA  
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT  
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

131/330

**FIGURE 131**

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 13000 14000 15000 16000 17000 18000 19000 20000 21000 22000 23000 24000 25000 26000 27000 28000 29000 30000 31000 32000 33000 34000 35000 36000 37000 38000 39000 40000 41000 42000 43000 44000 45000 46000 47000 48000 49000 50000 51000 52000 53000 54000 55000 56000 57000 58000 59000 60000 61000 62000 63000 64000 65000 66000 67000 68000 69000 70000 71000 72000 73000 74000 75000 76000 77000 78000 79000 80000 81000 82000 83000 84000 85000 86000 87000 88000 89000 90000 91000 92000 93000 94000 95000 96000 97000 98000 99000 100000

**FIGURE 132**

GGGGAATCTGCAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG  
GCTGCTGTTGTTCCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA  
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC  
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG  
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA  
GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGGTGGAGCACTTTATT  
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGTATCAATGTACGAGATTATCCTCA  
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC  
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT  
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG  
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG  
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC  
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT  
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA  
AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGTATGAGTGGCTAGAATTCTTC  
TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA  
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA  
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG  
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT  
TCCCAAAATGTTGAAAACCTGAACTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCA  
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA  
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA  
AAGATTTAAAATGTGTCTAATACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA  
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC  
ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA  
TGTGATGATGCCCTTTGTCCCATTTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA  
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG  
AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG  
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG  
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT  
TCAGGTTCCCTTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA



**FIGURE 133**

MEWWASSPLRLWLLLFLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCSSQNCSCYHGVIE  
EDLTPFRGGISRKMMAEVVRRLGTHYQITKNRLYRENDCMFPSRCSGVEHFILEVIGRLPD  
MEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDL  
FREDLVRSAQWPWKKNSTAYFRGSRTSPERDPLILLSRKNPKLVD AEYTKNQAWKSMKDT  
LGKPAAKDVHLVDHCKYKYLENFRGVAASFRFKHLFLCGSLVFHVGDEWLEFFYPQLKPWVH  
YIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMD DITCYWENLLSEYSKFLSY  
NVTRRKG YDQIIPKMLKTEL

**FIGURE 134**

CACCCCTCCATTTCTCGCC**ATG**GCCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT  
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCTGTGCGCTTTACCTCCCTTCGGCCACTTCTT  
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA  
CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC  
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCTTTCAG  
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC  
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC  
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT  
GACTATGCTGAGCTCATGGGCCTCAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC  
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG  
TGGAGCTGCTGACAGTGCTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT  
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT  
CCGGGCCCAGCTACAAAGAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAG**T**  
**G**AGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCTCCTCTCCCACTGAATTCTAAATCCTTAAC  
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT  
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTTCTGAGTTTCAGCCA  
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC  
TTCACTGTTTtagagcatgacactctccccctcaacagcctcctgagaaggaaaggatctgcc  
ctgaccactcccctggcactgttacttgccctctgCGCCTCAGGGGTCCCCTTCTGCACCGCT  
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA  
GGCCCCAACCTTGCCCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT  
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC  
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTTCTAAAAAAA

**FIGURE 135**

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP  
LAWDLGLLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV  
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP  
RALRLFSHLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR  
KLHLLSRPQDGEAE

**Signal sequence:**

amino acids 1-13

**Transmembrane domains:**

amino acids 58-76, 99-113, 141-159, 203-222

**N-myristoylation sites:**

amino acids 37-43, 42-48, 229-235

**FIGURE 136**

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA  
 AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT  
 TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTTATCTGGAATT  
 CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA  
 TCACGTGGTGTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG  
 AGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTAGGATGTCACTGAGATCC  
 CTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT  
 TCCCCACTACAATGTGATAGAACGCGTGAACCTGGATGTACTTCTATGAGTATGAGCCGATTT  
 ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACCTGCTCTCATCAAAATCCATTT  
 CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC  
 TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG  
 AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT  
 GACATAATCCGACAAGATTTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC  
 ATTCAGGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG  
 TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT  
 TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT  
 TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATAA  
 TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTT  
 GAAGATGTTTATGTGCGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA  
 CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG  
 CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC  
 ACATGCCATTATTAACCTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA  
 GTGTTAAATAAAGTAGGTACTGTGGAATAATCATGGGGAGGTCAGTGTGCTGGCTTACACTG  
 AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC  
 AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA  
 GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG  
 TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTTATTG  
 AACAAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA  
 TGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT  
 TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTATTTAAAATTA  
 CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT  
 CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC  
 TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTGGACTTTGTTAAAT  
 ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

**FIGURE 137**

MASALWTVLPSRMSLRSLKWSLLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQD  
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EK  
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTD TDVFIN  
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD  
LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG  
FSSKEIITFWQVMLRNTTCHY

**FIGURE 138**

CCTCTGTCCACTGCTTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT  
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA  
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA  
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC  
AGACTCTTTCAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT  
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC  
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA  
AAAAACATTGCAAACATGTGTTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC  
AAGCCTGTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT  
CCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT  
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT  
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

## **FIGURE 139**

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNGWDSWNS  
IWDYGNNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGGPPPKGLMYSVN  
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

**Signal Peptide:**

amino acids 1-20

**N-myristoylation Sites:**

amino acids 67-72, 118-123, 163-168

**Flavodoxin protein homology:**

amino acids 156-174

[illegible][illegible]



**FIGURE 141**

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLFLFGTTEEEIQEICIETLRLY  
TRKKPNYELLEKEVEKRRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK  
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN  
NRRSRSGTYSSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRRKRSRSQ  
SKSRDHSDAAKKRHERGHRDRRERSRSFERSHKS KHHGGSRSRGHGRHRR

**FIGURE 142**

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA  
TTTTTTGAANNNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT  
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT  
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNNTTATTGCTTACTGATTTTTTTG  
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAAGA  
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGCTGGTTTGTTCATGCA  
AGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT  
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA  
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC  
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTACTG  
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT  
GAATCCGGATGGAAGCTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

**FIGURE 143**

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTCA**ATGGC**  
GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC  
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTTTGTTCGGATAACAAGCACCCG  
TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG  
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG  
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG  
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT  
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT  
CCATCAAGATGTTCTTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC  
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG  
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC  
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG  
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC  
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCTCCTGC  
ACACCAGCTTCCTGTCTCCCCTGTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC  
TTCTTGCAACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA  
CTCTGGGCGCCTCTGGTTGCTGGTGGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC  
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC  
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT  
GAGCTTGACGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC  
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC  
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG  
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT  
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA  
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA  
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC  
GGCGTTCCCTTCACCACAGTGCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA  
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCAGGCCATTGTCTCCGAAGCGTATG  
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC  
GATTTTTTAA

**FIGURE 144**

MAVLGVQLVVTTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEEELRALAGKPRPRGRKE  
RWANGLSEEKPLSVPRDAPFQLETCPLETTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY  
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV  
VREETLELGLEPGLASMTQNLEPLLKKQGDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR  
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA  
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT  
VVSLQYLTPILTLNCTLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI  
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLGS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

145/330

## **FIGURE 145**

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC  
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT  
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC  
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC  
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC  
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT  
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT  
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTTCCTGACAGTGACACGGCTGTACTTCAGC  
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC  
CATGCTGGTGCAAGCG

**FIGURE 146**

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGGCCCGTGATTTATTAAACGTGGCTT  
AATCTGAAGGTTCTCAGTCAAATTTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGC  
TTGGCTGGTTTTGGGCCCTTGTAGCTGACAGAAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGG  
CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG  
TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG  
TCCAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA  
TGACAGACGAGCCTGGCCTAGACAACCCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCCA  
GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTAA  
AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG  
GCAGGGAAAAATTCTGAAACACCACTGCCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA  
ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC  
CCCACTGGTCCATATCATTTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG  
GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGG  
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCGCAGCAGGAACAATGGACAGGCCCC  
GGATGCCCTACAGACCCCGAGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGAA  
TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA  
CATGGTCAGCTTGAGGAGAATGACCGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAG  
TGCGGCTCATCTGATTTCAGGCCAGTGAAAGACGTGTTACCTCGTGTGTCGCGCCAGGTTCCGGCAGCGGAGCC  
CTGACATCTTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT  
CCCAAGCCCCCTCCATCCTACAATTACTTGTATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT  
CGGCATGACCGTTCGCGAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAGCCCG  
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTACA  
GAGGTCAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT  
CAAAGATATGAGCCCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCA  
GTGACTGGTCCCCATCCTGGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA  
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTT  
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG  
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGCTGGCAAGACTGCTGAAAGAACTTAAAGGAAGA  
ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTATAGAAATCAATGATGGGTGAGGAAAAACAGAAAAA  
TCACAAATAGGCTAAGAAGTTGAAACACTATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGT  
AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA  
AACTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTCTATTCAAT  
AAAAAGCCCTAAAACAACATAAATGATTGATTTGTATACCCCACTGAATTCAGCTGATTTAAATTTAAATTT  
GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCTAAAATATTTTTTAAATGCA  
TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAAGTTAAA

**FIGURE 147**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT  
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS  
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP  
SESLSIRLVGGSETPLVHIIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL  
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV  
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVSRQVRQRS  
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE  
WDLPIYVISVEPPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV  
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV  
GGYEEYNGNKPFIFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI  
TLTIVSWPGTFL

**FIGURE 148**

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT  
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT  
TTGGGATACAATCTCATGGATACGAGGTTTTTAAACATCATCAGCCCAAGCAACAATGGTGGC  
AATGTTCAAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC  
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC  
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT  
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG  
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC  
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT  
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG  
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC  
TTGGTTTACACTCAAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC  
AGTCAGAAAATATAAATGCTGTATTTATA



149/330

**FIGURE 149**

MKILVAFLVVLTIFGIQSHGYEVENIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT  
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNQLQWYIYEKQALDNMFSNKYTWVKYNPLE  
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 150**

GGCACGAGCCAGGAAGTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC  
**ATG**GGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG  
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG  
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC  
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT  
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCTT  
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG  
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA  
GATGATCTGCCAGGCGTCCTCGGGCAGCCACCTATCACCAACAGCCTGATCGGGAAGGATG  
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCTGCCG  
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC  
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA  
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG  
GGGTTTCAAGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC  
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA  
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

Downloaded from www.elsevier.com

151/330

## **FIGURE 151**

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL  
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK  
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP  
SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG  
GFRIGNGEVRGRKAAAM

### **Signal Peptide:**

amino acids 1-18

### **N-glycosylation Sites:**

amino acids 86-89, 132-135, 181-184

**FIGURE 152**

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG  
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT  
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA  
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA  
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT  
TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGC  
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT  
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC  
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTGTGGCCATGTCCTTCCATT  
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC  
CTGGAGCCAAGTGCAGGAGCACCCTCGCCATGTCTCAGGCACAACCCAACCTCAGGGCCAC  
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA  
TCTGAGGAGAGTCCTTTAGAGTGACAGGTAAAGCTGATACCAAAGGCTCCTGTGAGCACG  
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT  
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT  
TTGCCAACAAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC  
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC  
TCTTTTTGTTTGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT  
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG  
AATTTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT  
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

**FIGURE 153**

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL  
HYDCGNKTVTPVSPPLGKKLNVTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR  
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS  
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

**Important features:****Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 224-246

**N-glycosylation site.**

amino acids 68-72, 82-86

**N-myristoylation site.**

amino acids 200-206, 210-216

**Amidation site.**

amino acids 77-81

**FIGURE 154**

GGGAAAGCCATTTTCGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG  
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGA**ATG**GAG  
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG  
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAATCATGAACTCTGGCTGTTGATCAAAA  
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAGAAGCTAGCAGAAGAC  
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACAT  
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAACTCAAATTGGGAGGCCAAC  
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC  
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT  
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC  
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTCTGTAT  
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

155/330

**FIGURE 155**

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKLA  
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

155/330

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATTGGTGAATTATGAACGTGCCACACC**ATGAAG**  
CTCTTGTGGCAGGTAAGTGTGCACACACACCTGGAATTGCCATCCTGCTCCCGTTCTGCTA  
CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCTCAGCCGGGCCCC  
AGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAAGTTCAGCAAGGTGGTGTGCACGCGCCGG  
GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA  
CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCACT  
TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCTTCAACGGCCTGGCCAGCCTCAAC  
ACCCTGGAGCTGTTGCACAACCTGGCTGACAGTCATCCCTAGCGGGGCTTTGAATACCTGTC  
CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA  
ACCGGGTGCCCTCCCTCATGCGCTTGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT  
GAGGGAGCTTTTGAGGGCTGTTC AACCTCAAGTATCTGAACCTGGGCATGTGCAACATTAA  
AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT  
TCCCTGAGATCAGGCCTGGCTCCTTCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTGATG  
AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAC  
CAACTTGGCCCAATAACCTCTCTTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC  
TGGTGGAGTTGCATCTACACACAACCCCTTGGAAGTGTGATTGTGACATTCTGTGGCTAGCC  
TGGTGGCTTCGAGAGTATATACCAACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT  
GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT  
TCATCATGGAGCGACCTCGAGACCTCAACATTTCTGAGGGTTCGGATGGCAGAACTTAAGTGT  
CGGACTCCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC  
CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAACTTTTCCACGTGCTGC  
TTTCAGACACTGGGGTGACACATGCATGGTGACCAATGTTGACAGGCAACTCCAACGCCTCG  
GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAGT  
AACAGTGGAGACCACGGAGATCTCGCTGAGGACACAACGCGAAAGTACAAGCCTGTTTCTTA  
CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC  
CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT  
GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG  
CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC  
ACAGCCGCCCCGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC  
AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACAAATTC  
ATGACCATATTAACCTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC  
CTGGGGAACCTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTAGACCCA  
TACCAAGGACAAGGTACAGGAACTCAAATA**TG**ACTCCCCTCCCCCAAAAACTTATAAAAT  
GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA  
TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGA  
CAAAAAGTCAAAACA



**FIGURE 157**

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT  
RRGLSEVPQGI PSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS  
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY  
ISEGAFEGFLNLYLNLMCNKDMPNLTPLVGLLEEMSGNHFPEIRPGSFHGLSSLKKLW  
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHPWNCDCDILW  
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL  
KCRTPPMSSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN  
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ  
TTRVPKQVAVPATDTTDKMQTSLDEVMTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQORS  
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE  
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

**FIGURE 158**

CGCTCGGGCACCAGCCGCGGCAAGG**ATG**GAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC  
TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATG  
TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCCTGCCCCGAAAGAGGGAAGTCGTGGGTTATACCAT  
CCCTTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAACT  
GCAAGAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGATGACTTCTATGTGAAGGGGTCTACTGTGCAGAG  
TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT  
GTTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTATGCTAAACCTGGGTTTGTATCCAATAA  
GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCCGTGATGGAGACAAC  
CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACT  
CCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT  
GCTCCTCATCCCTTGTTCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG  
GCAGGCTATACTGGGCAGCGCTGTGAAAATCTCCTTGAGAAAGAAAAGTCTCAGACCCTGGGGGGCCAGTCAA  
TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTTGGCACCGTGGTGTCTT  
TCTTTTGTAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAAGTGGCCAGCAGAATGGAGAGTGGTCAGGG  
AAACAGCCCATCTGCATAAAAAGCCTGCCGAGAACCAGGATTTTCAAGCCTGGTGAGAAGGAGAGTTCTTCCGAT  
GCAGGTTCACTCAAGGGAGACACCATTACACCAGCTATACTCAGCGGCCTTCAGCAAGCAGAACTGCAGAGTG  
CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGGATACCAACATCTGCATACCCAGCTCCAG  
TATGAGTGCATCTCACCCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG  
TGGGCGGGCACCATCCTGCATCCCTATCTGCGGAAAATTTGAGAACATCACTGCTCCAAAGACCCAAAGGTTGC  
GCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG  
TTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGG  
GAAGGTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAAATTTACCGGGATGATGACCGGGATG  
AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCT  
GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCCTCGCTGCCAG  
TCGGGATCTCAGCACTTCCCTCCAGGAGTCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGA  
GCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGTGCTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG  
CATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCTTC  
TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGGACGAGCATCTCCTGAGCCACGCT  
GGCATCTGATGGGACTGGTCACTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTCACCAAG  
GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAAT**TGA**ACCATGCTCATGCACTCCTTGAGAAAGTGTTC  
TGTATATCCGTCTGTACGTGTGTATTGCGTGAAGCAGTGTGGGCCTGAAGTGTGATTGTGGCTGTGAAGTGG  
CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT  
GCCGCGTCCACTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACC  
ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTTCACTTATACGAATGCCATCAGCTTG  
ACCAGGGAAGATCTGGGCTTCATGAGGCCCTTTTGGAGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC  
CAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTTTGTGTACATGGCCACAGTACAGTCTGGTCCCTTTTCTTCC  
CCATCTCTTGTACACATTTTAAATAAAATAAGGGTTGGCTTCTGAAGTACAAAAAAAAAAAAAAAAAAAAAA  
AA  
AA

**FIGURE 159**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVV  
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGD  
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD  
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFAIYEEITACSSSPCFHDG  
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV  
VSFFCNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH  
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK  
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE  
RTVVVAAHCVTDLGKVTMIKTADLKVVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD  
ADIAILKLLDKARISTRVQPICLAASRDLSTSFAQESHITVAGWNVLADVRSPGFKNDTLRSG  
VVSVDSELLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR  
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

**FIGURE 160**

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA  
AGCTTTTCTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAACATGGGC  
TTCAACCTGACTTTCCACCTTTCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT  
GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAG  
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT  
AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAG  
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC  
CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC  
CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT  
CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT  
TTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC  
TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA  
GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG  
GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTTCAAGGTGAATGGATTCTCT  
AACAACTACTGGGGATGGGGAGGCCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG  
AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG  
ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGG  
AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA  
TATCAACATCACAGTGGATTTCTGGTTTGGTGCATTGACCCTGGATCTTTTGGTGATGTTTGG  
AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA  
AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTTCTTTTTGTATTTTCTTAGCAGAGCT  
CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG  
AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT  
AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT  
TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT  
CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCT  
GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA  
GGACACAGTGAACCTTGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGCGGCTGCAAAGGCAG  
CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT  
TCCAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTTAAAGAGTTTTTTGTAAATGA  
TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA  
TGTCTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

## **FIGURE 161**

MGFNLT FHLSYKFRL LLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT  
LTNEASTKKVELDNCPSVSPYLRGQSKLIEKPDLTLEEVQAENPKVSRGRYRPQECKALQRV  
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEEN  
WDCFI FHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG  
FSNNYWGWGGE DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR  
VWR TDGLSSCSYKLVSV EHNPLYINITVDFWFGA

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites:**

amino acids 4-7, 220-223 and 335-338

**Xylose isomerase proteins:**

amino acids 191-201

**FIGURE 162**

CGTGGGCCGGGGTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTC  
GAGCTCCAGCTGCATTCCTCCGCTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG**  
GCCCAGGCAGTGTGGTTCGCGCCTCGGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC  
CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA  
CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG  
CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGGCTTACTGGCAA  
GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTTCGGCCACGTGCCCCGGGAATTCCCGG  
TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG  
GTCCTCCCCATCACAGAGTTCTCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC  
CTGGCCCAGCTCCTATCTCACTAAGACCGTCTGAAAGTCTCCTTCCCTCCACGACCCGA  
GCAACTTCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG  
GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT  
CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA  
CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCTTCGAGGCATCCAAGTGTGGGG  
CCCACCCTAATTGAGACCTTCCAAAAGATGACCGTGACCTTGAACCTTCTGGGGAGCCCTCC  
TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC  
CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCTTGGGGACTAC  
TGCTTCAGCATCCGGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT  
GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA  
TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG  
AACCCGGAGCCACCCTCTGGGGTCAAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT  
GGAGACTCCATCTGAGTACCTGGAAATTGTTTCGTGAGAACCACGGGCTGCTCCCGCCCCTCT  
ATAAGTCTGTCAAACTTACACCGTG**TGA**GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA  
CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTTATT  
TGCGTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTTACGCCACTGCCACAAGCC  
CCTCCCTCTCTGTACCCCTGACCCCAAGCATTACCCATCTGTACAGTCCAGCCACTGACA  
TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT  
TTGATGCTTGGGGTGTTCGGTGTGACTCCTAGGTGGGCCTGGCTGCCCACTGCCCATTCCT  
CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCTCCCCCAGACAGCCCTAC  
CTGTGCCAGAGAGCTAGAAAGAAGGTATATAAGGGTTAAAAATCCATAACTAAAGGTTGTAC  
ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACA  
CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTTCAGATGATCAGCTCTGTA  
TCTGGTTAAGTCGGTTGCTGGGATGCACCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA  
AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC  
GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCTGGATGGGGGGCAGGACT  
AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG  
AACTTTTCACTGAGGAAAAGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG  
TGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA  
GATCGAGACCACCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAGTT  
AGCCGGGCGTGGTGGTGGGTGCCTGTAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG  
GTGCGAACCCGGGAGGCGGAGCTTGCAGTGAGCCAGATGGCGCCACTGCACTCCAGCCTGA  
GTGACAGAGCGAGACTCTGTCTCCA

163/330

## **FIGURE 163**

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA  
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF  
VVLPITEFLVGDLVVTQNTSLPWPSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ  
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL  
GPTLIQTFQKMTVTNLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD  
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAPFCATLITVMLAFIMYMTLRNATQQKDMV  
ENPEPPSGVRCCCQMC CGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 339-362

**N-glycosylation sites.**

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

**FIGURE 164**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCT  
GGGCCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCA  
CAACAGACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTG  
GATGCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCT  
GCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTG  
CCCCGTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAA  
TGGCTGGTTCTTTTGTTCCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AA



165/330

**FIGURE 165**

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPPMFQRRRRRDTH  
FPICIFCCGCCHRSKCGMCCKT

165/330

CTGTGACGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGAGGTTGCTGGCCTGGC  
CTGGATCTTCCACC**ATG**TTTCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC  
ATCTCCCTGACTGTCTCTTACCCTCCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTTGG  
AGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA  
CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC  
GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCTGTCGAAGTGG  
TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGGC  
GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG  
GAGTCTTGGAACTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC  
GGTCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC  
TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACCTGTGGTGGGATACTTGCCAAAT  
GGGAGGTTTAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG  
AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT  
GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC  
ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTACAGAGAGCCATGGTGAAGGCCTG  
CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA  
CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCAGAAGGAACCTGCATC  
AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTGAATTTGGAGCCACAGTTTACCC  
TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA  
TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG  
CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC  
CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCCTGAAGAGGGAGAAGG  
TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC  
AAGGACAGGAGCCGCTCCT**TGA**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA  
CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCCACTGCTGTGTCTTTTCCAGACTCCAGGG  
CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTTCGCCGAGCCGCAGCGGGATCCCTGT  
GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA  
CGAGATGCCTTGTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACCTCCCCA  
CCTTTGCACGCTGTGCGGGCTGAGTGGTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT  
GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG  
CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC  
CTTGGAGCTCTGCAGACATGATAGGAAGGAACTGTATCTGCAGGGGCTTTCAGCAAAATG  
AAGGGTTAGATTTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG  
GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTCTGTACTCCAGGCTCAACCTGAACTCCCC  
ATGTGATGCGCGCTTTGTTGAATGTGTGCTCGTTTCCCCATCTGTAATATGACTCGGGG  
GAATGGTGATGATTCCTACCTCACAGGCTGTGTGGGGATTAAAGTGCTGCGGGTGAGTGA  
AGGACAGATCAGCTTCAGTGTTCGAAGTACAGGCCACAAAACGGGGCACGGCAGGCCTGAG  
CTCAGAGCTGCTGCATGGGCTTTGGATTTGTTCTTGTGAGTAAATAAAACTGGCTGGTGAA  
TGA

167/330

**FIGURE 167**

MFLLLPFDSLIVNLLGISLTVLFTLLLVIIVPAIFGVSEFGIRKLYMKSLLKIFAWATLRME  
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSKSSKALDNTPEFELSDIFYFCRKGME  
TIMDDEVTKRFSAAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIIRYCFLPLRIALAFTG  
ISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRLTAIITYHDRENRPNGGICVANH  
TSPIDVIILASDGYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ  
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL  
LRMMTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF  
KEEQQKLYSKMIVGNHKDRSRS

167/330

**FIGURE 168**

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA  
CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC  
ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA  
CGATACCTATGTGGGTGTGTCAGAGGAGATGGAGGAGAAGGCAGCCCCCCTGCTAAAGGAGG  
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCCAGGAGACCTGGGAGGAC  
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT  
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG  
GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG  
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGG  
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT  
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT  
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCCTG  
GAAGACTCTGCTCTTGGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCTGAAAGTCCA  
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG  
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG  
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTATGTGATGGGGACTTCCT  
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTGCAATGTGGAGACA  
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

**FIGURE 169**

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH  
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSTLYWELNQAVRTGGGSREL  
YMRHFPPKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS  
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 170**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCTCAACATGCCTCA  
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG  
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC  
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCA  
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG  
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT  
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG  
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG  
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC  
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT  
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT  
GAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCT  
CCTGCTCAGTCTCTTTGTAAGTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG  
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGCGGAAACTCCTAACATATGCCCCCAT  
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA  
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAATCCCCACTCAC  
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGCAGTG  
CACTCCCCTAAGTCTCTGCTCA

171/330

**FIGURE 171**

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFENTTPLVT  
IQPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV  
YEHLSPKPKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW  
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW  
FLKRERQEEYIEKKRVDICRETPNICPHSGENTHEYDTIPHNTILKEDPANTVYSTVEIP  
KKMENPHSLLTMPDTPRLFAYENVI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 172**

CTGGTTCCCCAACATGCCTCACCCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC  
TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC  
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATAC  
AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA  
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT  
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG  
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG  
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT  
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG  
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC  
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCT  
GTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC  
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAA  
ACTCCTAACATATGCCCCCATCTGAGAGAGAACACAGAGTACGACACAATCCCTCACACTAA  
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA  
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG  
AATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA



**FIGURE 173**

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTGCT  
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA  
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA  
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGAGAGAA  
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG  
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT  
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA  
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA  
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT  
GCCAATCGCTGCCTCCAGGCCCAGAAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC  
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC  
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC  
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT  
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

174/330

**FIGURE 174**

MKMLLLLCGLTLVCVHAEESSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ  
IHVLENSLVLVKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI  
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

174/330

175/330

## **FIGURE 175**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA  
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG  
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC  
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCTTGACAGCAAG  
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA  
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC  
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA  
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG  
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC  
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT  
TGTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC  
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTA  
TCAGTAGTTTGAAAAAAAAAAAA

**FIGURE 176**

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA  
IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA  
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDFSEENKHRL  
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCT  
 GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT  
 GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC  
 CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCGACCAAGTTGGAAGTGAAGCACTG  
 CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATAGTGAA  
 AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC  
 CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

178/330

**FIGURE 178**

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV  
KHCTDQISFKKRLSLKKSWWK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 179**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG  
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG  
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC  
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG  
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG  
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG  
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC  
TCATCACTCCAGGCTCTGCCACTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA  
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC  
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG  
CCCCACCACCCCCTCA

180/330

**FIGURE 180**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK  
YKSSQKQHSPVPEKAIPLITPGSATTC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 181**

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC  
TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC  
CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC  
AGTGTTCCTCCACAGCCCCCAAACGGAACTGGTTTTTGGGGTCACCTGGGCCTGATCACTCCT  
ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT  
ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA  
CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCTGAAGCCCTGG  
CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC  
GCCCCGCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA  
TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTTGAG  
CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG  
TCAGGAGAGGGCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA  
GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC  
TCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCCG  
CACCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG  
ATTTCATTTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT  
ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC  
CTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC  
AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCC  
TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG  
ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC  
TCATCGATATTATAGGGGTCCATCACAAACCAACTGTGTGGCCGGATCCTGAGGTCTACGAC  
CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTC  
CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTGCCCATGGCGGAGATGAAAGTGGTCCTGG  
CGTTGATGCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAA  
TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCA  
**GTGA**CTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGTGTCAA

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPPKRNWFWG  
HLGLITPTEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF  
IRFLKPWLGEGILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGS  
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY  
LSHDGRRFHRACRLVHDFTDVIRERRRTLPTQGIDDFFKDKAKSKTLDLFDVLLLSKDEDG  
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW  
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW  
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT  
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

**FIGURE 183**

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC  
TGTTTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA  
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA  
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG  
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT  
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA  
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTCTGAAGTGAAATAAGCATCTGT  
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC  
AGTGTGGAGAAAACTAGGCAAACTACACCCTGTTTCATTGTTACCTGGAAAATAAATCCTCT  
ATGTTTTGCACAAAAAAAAAAAAAAAAA

184/330

**FIGURE 184**

MYKLASCCLLFTGFLNPLLSLPLLDSEISFQLSAPHEDARLTPEELERASLLQILPEMLGA  
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GAACATTTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT  
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA  
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCC  
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT  
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA  
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA  
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT  
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG  
GAAGAGGCCAAAGAGGCCCCAGCCGACAAGTGATCGCCACAAGCCTTACTCACCTCTCTCT  
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA  
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGTGTA

186/330

**FIGURE 186**

MPSPGTVCSELLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG  
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTTCCTCGAGTACCTC  
GTGAACCCCGGGGTGCTCCGCACGGACCCAGATGTCAAGAATATGAACACGTGGCTGCTGT  
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG  
CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT  
GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC  
TTACACCTCTACTTGAGTATGTCCCTAACCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT  
GTCCCCCGTGTGCGCATGTGTTCAAGGTGAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA  
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC  
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAACAAGTTTAAATGTTTACA  
GACAAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT  
CCAAAAACACAAGTAGAAATTCTAACAAATGAAATATATTACAGGCAGGTACCCACTAACCA  
AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC  
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT  
AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCG  
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC  
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA  
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGA  
GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCACCAAGA  
GCCTCCTTGTTTCATAACCACAGGTTACCCTACAAACCACTGTCCCCACACAACCTTGGGGAT  
GTTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA  
ATTTTTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA





**FIGURE 189**

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCCAAG  
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT  
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG  
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCTG  
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACCTGGGAGACTGGGGATGA  
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG  
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA  
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA  
GCGGTTGATGGAGAAGGCTTCCCTCCCCCTCCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA  
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTATCAGCTTCCTCCTGCT  
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACCTGAGCGCCTTTGCTG  
CTGTTTCCTCTGTCTCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTCACAAGTC  
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGAATTATGGCTG  
GGCCTTCTACATGGCCTGGCTCTCCTTACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA  
ACACGTACACCAGGATGGTGCTGGAGTTCAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC  
CCGAAC TGCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGACGCCCCAC  
CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG  
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGGCCAGCCAGGAGCTG  
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA  
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG  
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC  
CTAAGGGATTCTGGGTGCCACTGCTCTCTTTTCCCTCTACAGCTCCATCTTGTTTCACCCAC  
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTCTCTGGGC  
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA



**FIGURE 191**

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG  
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA  
GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG  
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT  
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG  
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCCCTCCTCCCAA  
GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG  
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT  
AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCGCCGTAGATTCTAG  
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT  
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT  
ATGGCGATGGCCATGATGTTACAATCCCCTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA  
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAA  
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC  
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA  
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAACTCTTTATTACTTTGGG  
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG  
AACCAAGGAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC  
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA



**FIGURE 193**

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCTGGAGCCAGGAGCGACGTCA  
 CCGCC**ATG**GCAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTT  
 GGATGTGCCCTTCCAATATACAACAAATACTGGCCCCCTCTTTGTTCTATTTTTTTTACATCCTTTCACCTATTCC  
 ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGCCATCTTTC  
 TTACAACGGGCATTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGA  
 GCTTGTGCACTTGTCTCACAGGAAACACAGTCATCTTGCACCTATACTAGGCTTTTTCTTGGTCTTTGGAAG  
 CAATGACGACTTCAGCTGGCAGCAGTGGTGA**AAA**AGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCAATTT  
 GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA  
 GGTGCTCCCTTCTCACTTTTATTGTAAAGCATACTATTTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT  
 CTTTTGGAAAAGCTTGACTGATTTTCACACTTATCTATAGTATGCTTTTTTGTGGTGTCTGTGAATTTAAATAT  
 TTATGTGTTTTTCTGTAGGTTGATTTTTTTTTTGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA  
 TTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACATCTTTTCTTTAAATATTATT  
 TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATG  
 CAGACATACAGACGGTTGGCATACTGTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG  
 GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTGTGGTGGTTTACTGGTAGACAGATGTTTTGTGGATTG  
 AAAATTATTTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA  
 AGGTAAGGGTGTAAAAACATTTTTGAGATAAGGTTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGT  
 GGAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC  
 AACCTTTCATGTTTTACCTGTATAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC  
 ATCATATATGCCAGAAAACCTTCTCTGCTTCTCTCTTTGACTTATTTGGTATGTTGTATATATTACATAAAA  
 TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT  
 CAGAGTGCCCCCTCCCCTGCAAGGCCCTGCCATGATTAACAAGTAACTTGTAGTCTTACAGATAATTCATGCA  
 TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAG  
 TATTTTTAAGACAAGTTTCTGTATACCTCTGAACGTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT  
 CCTTATAAAAGGCATTTGTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT  
 ACCTGACCAAAAAATTTCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA  
 CTTTTTTTCAGGAGTGGGTTATAAAAAACATTCAAGTTGGTCTGACAGTATTTTGTAAAGGATATTTGTTGTATG  
 TTTATTCAGTATACTTACATAAAAAATTATTTGCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTGT  
 TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTTCTGCTTTTACTAAAGATGCCTAA  
 AGCCACAGGTTTTATTGCCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG  
 GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAA  
 GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACCTCTTGTGCTTGTGATCTACTGGACTTT  
 TTTTTTGCAGGAAGTGCATTCTCTGGTCCTTCCCTATTTTTCTGTTCTGGATGTCAGTGCAGTGCAGTGTACTG  
 TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT  
 GTGTCTTTGACCTTGATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT  
 TTTCCATAGAATATGCACTGATACACATTACCATTCTTCTATGGAAAGAAAACCTTTTGATGATGAAACAATAA  
 AGATTTTTAAATATCTATTTTAAAAA



[illegible]

**FIGURE 196**

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH  
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ  
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS  
LKFLDIGYNQLKSLARNSFAGLFKLTLEHLEHNDLVKVNFAHFPRLLISLHSLCLRRNKVAIV  
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT  
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPTSG  
HLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA  
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH  
IEGALVIINEYGSCCTCHQQPARECEV



197/330

**FIGURE 197**

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCCTGCTGTGGCTGCAGCTCTGC  
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA  
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCCGTTGAGTTCCCGGCGGACAAGATGG  
TGTCAGTCCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA  
CTCGTCCTGGCTTCAGGAGCCGGATTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG  
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT  
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC  
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG  
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG  
CTGTTTTCTTGGCGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCTTGAGCGTG  
GGCCCCGAGGACTGCGCGGACCCGTCGGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG  
GATCTGCGCGGCCCTGCTCCAGCCCCT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

198/330

**FIGURE 198**

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE  
GHAUSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSRSGDEA  
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR  
AGRLRFHGPGLSVGPEDCADPSGCVCGNAEAQPWICAALLQP

198/330

**FIGURE 199**

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTT  
TGAGCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTT  
CTGCCTTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCA  
TTGACGTGGTACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTTTTGC  
CGTGTGTCTTGCATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAG  
CTGGTGGACAGTTTTGTAACCTCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCT  
TGCAGCAATGTGTTGCTTGTGATTGGAACATTTGAGGGTTACTTTTGAAGCAACAATACAT  
TCTCGAACCTGAATGTCAGTAGCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAA  
TCTTCCTCATGTACCTGTTTCCTCTCTGGATGTTGTCCCACTGAATTCCCATGAATACAAAC  
CTATTCAGCAACAGCAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

200/330

**FIGURE 200**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP  
FARDAVKKCFVCLA

200/330

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGGCGTTCCCTCCAGTACCCTCCCGCCGTTACCCGCGGCGCGC  
CCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT  
GGGAGAGAAAAATTAGGGGGGAGAAAGGACAGAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG  
AACTGATCAAGTACTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTTCATACTTGCTGCATGAGTCTTTT  
AACCACCTTTTCTCTCCAAGCTAGACAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT  
TATATAAAGTTCCAACGCCCCATTTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT  
TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTCAGAGAATCATGGGATTGTTG  
AAATGATATGTTTGATCCTATTCGGAACAAATCTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTT  
GGGAAGAAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA  
ACAGATGTAAAAATCATATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCAATTTGAAGATG  
AGTTGCCAAATTTGTTGAATGGTTTACGTCAAAGAGCCCAATAATCTTGGTCTTCTCTATTGGGAAGACCCCT  
ATGACATGGGCGACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCAATTCAGATATTGACAAGAAGTTA  
GGATATCTCATACAAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG  
AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCCTATACCCTGATTG  
ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAATTTTGATGAAGTCTATGAAGCACTAACTCACGCT  
CATCCTAATCTTACTGTTTACAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATGACAACAGTCGAATTC  
ACCAATCATAGCAATGGCTGATGAGAGGTGGCACAATTTTACAGAATAAGTCAGATGACTTCTGTGTAGGCAACC  
ACGGTACGATAATGCTGTAGCAGATATGCATCCAATTTTCTTTAGCCCATGGTCTCTGCCTTCAGAAAGAATTT  
TCAAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA  
CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA  
CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTATACCCTTATTTTCATAGGGGTC  
TCTCTTGGCAGCATTATAGTGATTGATTTTTTGTAAATTTTCATTAAAGCATTTAATTCAGAGTCAAAATGCTGC  
CTTACAAGATATGCATGCTGAAATAGCTCAACCATATTACAAGCCTAAATGTTACTTTGAGTGGAATTTGCCATA  
TTGAAGTGGAGATTCCATAATTTATGTCAGTGTTTAAAGGTTTCAAATTTCTGGGAAACCAGTTCCAAACATCTGC  
AGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACACGGACCAA  
ATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG  
ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACATGTAAAGTACCTTATAT  
ATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC  
TTGATTGAAATGACAACCTTTTGCACCCATGTACAGAATACTTGTTACGCATTTGTTCAAAGTGAAGGAAATTT  
TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA  
AAATTAATATGTGATAACCTTTGAACCTTTGAATTTTGGAGATGTATTTCCCAACAGCAGAATGCAACTGTGGGCAT  
TTCTTGCTTATTTCTTTCCAGAGAACGTGGTPTTCATTTATTTTTCCCTCAAAGAGAGTCAATACTGACAG  
ATTGCTTCTAAATATATTGTTTCTGTCTATAAAATTAATTGTGATTTCTGATGAGTCATATTACTGTGATTTTCA  
TAATAATGAAGACACCATTAATATCTTTCTTCTATATAGTTTCAGCAATGGCCTGAATAGAAGCAACCGGCA  
CCATCTCAGCAATGTTTTCTCTGTGTAATTAATTGCTCCTTTGAAAATTAATCACTATTAAATACATTAA  
AAATCAAAATTGGATAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 202**

MTSKFILVSFILAALSLSTTFSLQLDQQKVLVVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK  
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW  
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY  
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER  
LIELDQYLDKDHYTLDQSPVAAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN  
SRIQPIIAVADEGWHLQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD  
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF  
IGVSLGSIIVIVFFVIFIKHLSQIPALQDMHAEIAQPLLQA

**Signal Peptide:**

amino acids 1-22

**Transmembrane Domain:**

amino acids 429-452

**N-glycosylation sites:**

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

**Somatomedin B Domain:**

amino acids 69-85

**Sulfatase protein Region:**

amino acids 212-241

**FIGURE 203**

GGATTTTTGTGATCCGCGATTTCGCTCCCACGGGCGGGACCTTTGTAACTGCGGGAGGCCAG  
GACAGGCCCCACCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC  
AGAGAGGCCAAGCCCCCTTGCCCTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA  
CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC  
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA  
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG  
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC  
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT  
CAGCTCCCACAGGTTTTAGGTCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC  
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCCACAAGAATAACTATGCTGCCATG  
GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT  
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG  
TGGTCTCATTTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAAGTTTGAGGCTCTGGGC  
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT  
TAAGACACGTTTCAAGACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA  
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGAC**TGA**TGAGTTTGCTGTATC  
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT  
CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG  
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA  
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAAGAGAACAAAACC  
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPV  
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMRLKLFSSSHRFQVIIICLVVLDALLVLAELILDL  
KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSFTTSLRSWMPVVVVVVSFILD  
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS  
CSEKPLD



[illegible]

**FIGURE 206**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD  
 LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK  
 SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW  
 WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI  
 NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMAL  
 RKTGQYSGMLDCARRILAREGVAAFYKGYVPMNLGIIPYAGIDLAVYETLKNWLQHYAVNS  
 ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG  
 LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 284-304, 339-360, 376-394

**Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

**N-glycosylation site.**

amino acids 129-133, 169-173

**Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140



**FIGURE 208**

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI  
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD  
AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD  
QGANFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTSESEIKRR  
SHLQLLSKASLCVSSFFAISWALLPLSPYLMLK

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332000>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 209**

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT  
GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC  
AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG  
AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC  
CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACCTCTCTGGCTATCATCTTCA  
CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC  
GAGTGCGGGACGCAGCTCTGTGGATCTCGACATACCTGTGTTAGTTCCCTTCCCAGAACCCAT  
CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT  
GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG  
GTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT  
GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT  
GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC  
AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT  
GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT  
CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG  
AGTTTTGTAACTTCAAGTGCTGTTCACTGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA  
CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA  
GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT  
CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG  
GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA  
GTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCTTGCCCTGCCCCACCCATGAGGTAGGCAG  
AAATCCTCACTGCCAGCCCCCTTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCCACCTGAC  
TCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAAAAA  
AAA

210/330

**FIGURE 210**

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMPFL  
LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS  
RHTCVSSFPEPISP EWVRTRPFPILPFPPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CTTCTGTAGGACAGTCAACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG  
GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAAACT  
TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT  
TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCA**ATGA**AGGGGCATCCTCG  
TTGCTGGTATCACTGCAGTGCTTGTTCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT  
TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG  
CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT  
TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT  
GCTGAAGAACACTTTTCATTTTGTAAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG  
CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG  
AATCTAATGGAACCTCCTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTTGC  
TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTT  
CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGAGGAGTCA  
TCTTTTGAAAGTTTGAGTGTGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT  
TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG  
GGGACTGCTGCCCT**TGA**AGGTCCTGGGGCTGCACTTTGCCAGCACCCCATTTCTGCTTCTCTG  
AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC  
CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA  
AAGCACTGGTTTCATTCACTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

212/330

**FIGURE 212**

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR  
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE  
CPACYESNGTSCRGKPKWCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK  
TLGGVIFRKFEKANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 213**

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA  
GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCC**ATGG**TCCCCGCCGCCG  
GCGCGCTGCTGTGGGTCCCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG  
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG  
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA  
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC  
ACGGTGTCCACCGGCTTTAGCCGGTTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA  
GGGGTTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA  
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG  
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC  
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTACCCCTACCCACAG  
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC  
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT  
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC  
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT  
GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG  
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTGGAAACGGGTCA  
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA  
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT  
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAACAAAACCTGGAAAA  
CACAA

**FIGURE 214**

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI  
ILEDENDAMADADRLAGPAAAELLAATVSTGFSSRSAINEDGSSEEGVVINAGKDSTSREL  
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP  
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC  
TYQQCPCNRLREECPLDTSLCDTNCASQSTTSTRTTTTFPPTIHLRSSPSLPPASPCPALA  
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CCCGGGTTCGACCCACGCGTTCGGGGGAGAAAGG**ATG**GCCGGCCCTGGCGGCGCGGTTGGTCTGTGCTAGCTGGGGCA  
CGCGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTAACCGGACTGCGTACTGCAGTGCGAAGAGCA  
GAACTGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCC GCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCT  
GTCGGGACGACTGTAAGTATGAGTGATGTGGGTCACCGTCTGGGCTACCTCCAGGAAGGTACAAAGTGCCCT  
CAATTCCATGGCAAGTGGCCCTTCTCCCGTTCTGTCTTTCAAGAGCCGGCATCGGCCGCGCTCGTTTCT  
CAATGGCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCATGTACCA  
CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC  
CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC  
CGTGGGGCTGCGACCCAGCTGTGGTCAGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCT  
CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAAACGTGGCTATTGGCCTGGTCAACGTG  
GTGTGGTGGCTGGCCTGGTGCCTGTGGAACGACGGCGGCTGCCACGTGCGCAAGTGCCTGGTGGTGTCTT  
GCTGTCGACAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCACCGCTCTTCTGGGTCTGGATGCCCATGCCA  
TCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG  
AAGGAATCAGAGGACAAGTTCAAGCTGGACT**TGA**AGACCTTGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCC  
GCCCTGCTGGCCTCCCTTCTCCCTCAACCCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACTTGGACATGA  
AGGATGTGGGCCCAGAATCATGTGGCCAGCCCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG  
AAGGCCCTCCAGCATCTGGGACTCGAGAGTGGGCGAGCCCCCTACCTCTCTGGAGCTGAACTTGGGTGGAACTGA  
GTGTGTTCTTAGCTCTACCGGAGGACAGCTGCCTGTTTCTCCCCACAGCCTCCTCCCCATGCCAGCTG  
CCTGGCTGGGTCTGAAAGCCTCTGTCTACTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC  
CTTCTGTTACCAACCCCCACCCCTCCTCCAGGACACCACTAGGTGGTGTGATGCTTGTCTTTGGCCAGCCAA  
GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTACCCCTGACC  
GTTGCCCTAGCCAGGTTCCAGGAGGCCTCACCATACTCCCTTTAGGGCCAGGGCTCCAGCAAGCCCCAGGGCA  
AGGATCCTGTGCTGCTGTCTGGTTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCGAGGCTGAGTGCAATAGG  
TGACAGGGCCGTGAGCATGGCCCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTGTG  
CGGGGAAGAGGCTGTGGCTTCAAAGTGTGTGTGTGACAGGGGGTGGGTGTGTTAGCGTGGGTAGGGGAACGTGTG  
TGCGCGTGTGTTGGGCATGTGAGATGAGTGACTGCCGGTGAATGTGTCCACAGTTGAGAGTTGGAGCAGGAT  
GAGGGAATCCTGTACCATCAATAATCACTTGTGGAGCGCCAGCTCTGCCAAGACGCCACCTGGGCGGACAGC  
CAGGAGCTCTCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCCTTGGCCCGCTCCTGCAAA  
CTCACAGGGTCCCCACACAACAGTGCCCTCCAGAAGCAGCCCCCTCGGAGGCAGGAGGAATAAGGATGGC  
TGGGCTCTCTCATCCTCTTTCTCCTTGCCTTCGCATGGCTGGCCTTCCCTCAAACCTCCATTCCCCT  
GCTGCCAGCCCCCTTGCCATAGCTGATTTTGGGGAGGAGGAAGGGGCGATTTGAGGGAGAAGGGGAGAAAGCT  
TATGGCTGGGTCTGGTTTCTTCCCTTCCAGAGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGCAGGGGCC  
ACACTATGCCTGTGCCCTGGTAAAGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCTGCCCCACAGG  
AATAGAATGGAGGGAGCTCCAGAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTG  
CTCTGCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCAAACCTCAGGGACTCGGGT  
GCCTCGCCCTAGCTTCTTTGATACTGAAACCTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAACTCAA  
TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

216/330

## **FIGURE 216**

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW  
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSSRFLFFQEPASAVASFLNGLASLVMLCR  
YRTFVPASSPMYHTCVAFAVWSLNAFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR  
TVGLQHPAVVSAFRALLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR  
RLPHVRKCVVVVLLQGLSLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFSFLEDDSLYLL  
KESEDKFKLD

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domains:**

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

#### **N-glycosylation site.**

amino acids 40-44

#### **N-myristoylation site.**

amino acids 43-49

#### **CUB domain proteins profile.**

amino acids 285-302

#### **Amiloride-sensitive sodium channels proteins.**

amino acids 162-186

**FIGURE 217**

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA  
CTATGGCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTG  
CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC  
CAACCCAGAGAAGAGCAGCACCAAAGAAACGGAGAGAAAAGAAACCAAAGCCGAGGAGGAGC  
TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG  
CAGGCTGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCAAA  
ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA  
ACACCTACACATCTCAGGATCTCAAGAGTGCCTGGCAAAATTCAAGGAGGGGGCAGAGATG  
GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA  
GGAAGTGAAGAAAGACTTTGATGAGCTGAATGTTGTCAATTGAGACTGACATGCAGATCATGG  
TACGGCTGATCAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAAGAGAAGATTGCTGCGCTC  
TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTTCTTTGGTGG  
TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGT  
TTGTGCTGGGCGCTGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA  
GCCCTGCAGAAGCTGCTGGTCACTCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGGT  
CCTGTTTGCCTGTGCTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTTCTGAAGC  
TCGGGGGGCTGCAGGTCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTG  
CGCGTGGTCACACTGCTCTACGACCTGGTCAAGGAGAAGATGTTCCGCCGAGGAGGAGGCTGA  
GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG  
GCCTGTGGGAACAGGGCTGGTGCAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGAT  
GCCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG  
TCAGGACCCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCA  
GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC  
AGCTTGCTGAAGGAGCTGAGATTGAGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGA  
GGCTGAGGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT  
TGCCCATTAATGGAAACCTGAAGGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AA

**FIGURE 218**

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL  
DAEVLEVFPHTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLK GKRLDINTN  
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV  
RLINKFNSSSSSLEEKIAALFDLEYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF  
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL  
GGLQVLRITLVQEKGTEVLAVRVVTLTYDLVTEKMFEEEEAELTQEMSPEKLQQYRQVHLLPG  
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS  
LELQDGEDEGYFQELLGSVNSLLKELR

**Important features:****Signal peptide:**

amino acids 1-29

**Hypothetical YJL126w/YLR351c/yhcX family protein.**

amino acids 364-373

**N-glycosylation site.**

amino acids 193-197, 236-240

**N-myristoylation site.**

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

**Homologous region SLS1 protein.**

amino acids 68-340

**FIGURE 219**

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCC  
 CTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTTCGCGTTGCCACCCCA  
 CGCGGACTCCCCAGCTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCC  
 TTCCACCTGACCAGCC**ATG**GGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTGCGGTTTCGGC  
 CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT  
 CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG  
 TCCATGTGACCGACCGGTGAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCT  
 GTCTCTGTCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA  
 TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT  
 ATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT  
 GATGCACTTGGGCCAGGTGTGGTGGGATCCATGGAGACTCACCTATTACTTCTGACTTC  
 AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG  
 CCTGTGAGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTGGGAGTCACTACTGACATCG  
 GGACTGACATTCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT  
 TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTAGCGCAGCC  
 TCTTGTGTAAGGACT**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG  
 CCCATGACTGAGCCCAGCCCCAGCCCGGGTCCATTGCCACATTCTCTGTCTCCTTCTCGTC  
 GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTT  
 TACCAGGAGCAGCCTGGGTTTCCAGCCAGTCACTGACTGGTGGGTTTGAATCTGCACTTATCCC  
 CACCACCTGGGGACCCCTTGTGTGTCCAGGACTCCCCCTGTGTGAGTCTCTGCTCTCAC  
 CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT  
 GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC  
 CTGGGGTGCCCTCCTGATGTCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG  
 GTTGCCAAGAAAAGGGACCTAGTTTTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC  
 AGGCCTGAGGGGGAACCATTTTTTGGTGTGATAAATACCCTAACTGCCTTTTTTTCTTTTTT  
 GAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTCTTCTAACCTCCTTGGGCTATATTTCTC  
 TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT  
 GGGGGAAAGGAAGGAAGTGCATGTTTGGGAACCTGGCATTACTGGAACATAATGGTTTTAACCT  
 CCTTAACCACCAGCATCCCTCCTCTCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC  
 CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGGAAGGAGG  
 GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGT  
 ATCATTTTCTGCTGAGGGTGGAGTGTCCATCCTTTTAATCAAGTGATTGTGATTTTGACT  
 AATAAAAAAGAATTTGTAAAAA  
 AA

**FIGURE 220**

MGAAVFFGCTFVAFGPAFALEFLITVAGDPLRVIIILVAGAFFWLVSLLLLASVVWFILVHVTDR  
SDARLQYGLLIIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS  
FGIISGVFSVINILADALGPGVVGIIHGDSPIYYFLTSAFLTAAIILLHTFWGVVFFDACERRR  
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 221**

AAGCTGGTTTAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA  
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT  
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC  
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

09997443094  
4004 46350

GACCGACCGTTTCAGATGCCCCGGTTCCAGTACGGCTTCCTGATTTTTTGGTGCTGCTGTNTCTG  
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG  
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN  
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTGGNTGATGCAN  
TTGGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCATTAAATTCCTGAATTCAGCCTTT  
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA  
GAGGAG

**FIGURE 223**

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCTTTCCCTTTCCCCG  
GGGTCTGGGGTGACATTGCACGGGGCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC  
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACNTG  
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTGCGGTTTCGGCCCGGCCTTCG  
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA  
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC  
CGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTCTGTCC  
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA  
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG  
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG  
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

**FIGURE 224**

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG  
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC  
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCTGGTCAGGCCCCCACCCTTCCCACCTGA  
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC  
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC  
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA  
CCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC  
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT  
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG  
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT  
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

**FIGURE 225**

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCCAGGAGGAGGCAG  
 TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC  
 TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG  
 TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT  
 AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCCGT  
 CCTTGTGGTTTCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATGG**CTACAGCAAGACCCC  
 CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT  
 CTCGCCAACAATGATGTTTCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA  
 GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTTCGGATGACAGCAGCAGCCGCATCATCA  
 ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC  
 CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG  
 GAAGAAAGTTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC  
 AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCCTGGCTACTCCCACCCCTGGCCACTCT  
 AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTCGTCCCCTAAAGATGTCAGACCCAT  
 CAACGTCTCCTCTCATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA  
 CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCTCCAGTGCTTGAATATCAGCGTGCTAAGT  
 CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA  
 CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC  
 TGCAGGGACTCGTGTCTTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC  
 ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGAG**TCAT  
 CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG  
 ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT  
 CCTGGACTCAGGGTCTGCTTCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG  
 GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT  
 CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA  
 CTGAGAAGTGGAATAAAAAA

226/330

**FIGURE 226**

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD  
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS  
PVYESGQQMFQGVKSIHPHGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCL  
VSGWGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGRDSCQGD SGGP  
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.26.008650>; this version posted March 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 228**

MVNDRWKTMGGAAQLEDPRDPKQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT  
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA  
QPRLVGDQEQELLDTLADQLPRLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL  
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL  
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT  
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAARYGSFGVGLFSVDPEEDGYPLTVADYS  
GTAGDSLLKHSGMRFTTKDRDSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLGAHASADG  
VEWSSWTGWQYSLKFSEMKIRPVREDR



**FIGURE 229**

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT  
 TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT  
 CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG  
 ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCC  
 CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC  
 TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACC  
 AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC  
 CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC  
 TGAAAAACTCTGTCTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG  
 AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG  
 GACTGTAAATATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAACAAGAAGA  
 CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT  
 TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTCACTTCTGAACTG  
 TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG  
 GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA  
 TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGAC**TGA**TTCGCC  
 CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT  
 TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG  
 TTTCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA  
 AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCATAATGTCATCTGCCTTCTTG  
 GCTTAGAGATAACTTTTAGCTCTCTTCTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT  
 GTCTTCCTTACACTTGGTGGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC  
 ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC  
 AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT  
 CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG  
 AATCTCAAATCTCAATGCCTTATAAGCATTCTTCTGTGTCCATTAAGACTCTGATAATTG  
 TCTCCCCTCCATAGGAATTTCTCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG  
 AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA  
 TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG  
 GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

230/330

**FIGURE 230**

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL  
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE  
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS  
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD  
CKELKRCVCERRAGMVKPESLHVPPETLGEGD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 231**

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG  
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG  
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTTCAACGTGGCGACCAGTGGCCCTGACCCTG  
CTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTA  
CTACCAGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATA  
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT  
GTGGCTGAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGGAACTTTGAAGGAGGGCAA  
AGTNTCCTCATNTACTATACACACACCACTTCCC

CCGCGAGCGCAAGAACCCTGCGCAGCCCCAGAGCAGCTGCTGGAGGGGGAATCGAGGCGCGGCTC  
CGGGGATTTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCGCGCGGGG  
CCCGAGCCCTCCGGATCCGCCCCCTCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT  
CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCC**ATG**CGGGCATCGCTGCTGCTG  
TCGGTGCTGCGGCCCCGAGGGCCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT  
GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGCCAGGCCCGCCCCAACCTGGAGACTCTG  
AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGCCCCAACCTCGGTGCAGCCCGGAGCG  
GAGCGCGAGAAGCCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC  
CTACCACCCTGCACAGCCCCGGCCAGGCCGCCAAAAAGGCCGTGAGGACCCGCTACATCAGC  
CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCC  
ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC  
GGGCGCACGGGGCCGCGGGGCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC  
CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC  
TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG  
CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCGGGCCCCAGGACTTCATCGGCG  
GAGAGCCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTCGCGCATGCTG  
CTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA  
CGAGTGGCTGGGTGCTGCAATCTCGATGCCACCGGGGTGGGCTGCATGGTGACCACGAGG  
GGGTGCACATATAGCCATCTGGAGCTGAGCCCTGGGAGGCCAGTGCAAGGAGGGGACCCTCAT  
TTCCGAAGTGCCCTGACAGCCCCACCCTGTGCGTGACCTGTGCACATGTACCAGCTGCACAA  
AGCTTTCGCCCGAGCTGAACCTGGAACGCGACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA  
TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT  
CCAGCACCATCCCGCCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA  
GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCACTGCGTGGGGCTGACCGGGCTG  
ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCJGGCCTTG  
CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGGTATGGA  
ATACACGCTGGACTTGACAGCTGGAGGCACTGACCCCCCAGGGAGGCCGCGGCCCTCACTC  
GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCCTGTGCCCTATGTCACT  
GAGGCCTCAGCTCTCACTGTGCTGCTGCCCTTAGCTGCGGCTGAGCGTGACCTGAGCCCTGG  
CTTCTTGAGAGGCCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC  
TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT  
GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCGGTGCCCGGGTGCCATGGCTCAG  
TGTGCAGACAGCCGCACCCTCACCCTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC  
TGGACACACTGTTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC  
TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA  
CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGT  
TTGATCGCCAGGCAGCCAGAGGCCCTGCTTCTACAACCTCCGACTACGTGGCAGCCGCTGG  
CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGATGTGTACGAGCT  
GTTCTTCCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT  
ACCGGGCCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCCTCCAGAGC  
GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG  
CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCCACCCACTT  
CTCCCCCAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC  
AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG  
GGACGTGCCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT  
GCTGATTTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC  
CTCTGGGCCCTGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG  
GCATCTCCCAACTTCTCCCTTTTGACCCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA  
AGTGTGGAAAAA

**FIGURE 233**

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP  
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAACKAVRTRYISTELGIRQRLLVAVL  
TSQTTLPTLGAVVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLL  
QHGDDEFDWWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG  
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCIL DATGVGCTGDHEGVHYSHLELSPGEP  
VQEGDPHF RSALTAHPVRDPVHMYQLHKAFARAELE RTYQEIQELQWEIQNTSHLAVDGDRA  
AAWPVGIPAPSRPASRF EVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN  
RRYHPALRLQKQQLVNGYRRFDPARGMEYTL DLQLEALTPQGGRRPLTRRVQLLRPLSRVEI  
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLLYEPRQAQRVA  
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL  
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS  
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLR AVEPALLQRYRAQTCSARLSEDL  
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

**FIGURE 234**

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT  
 TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT  
 AGCTATTAGCCAATTTCGGCAGGGCCCCGCTTTTTAGAAAGCTTGATTTCCCTTTGAAGATGAAAG  
 ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA  
 CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA  
 GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG  
 AGGGTTGCTCAACGCCCCGCTCATTTGAAACCAAATCAGATCTGGGACCTATATAGCGTG  
 GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT  
 TTCCCCGCCCCCTGAGACCCTGCAGCACCATCTGTC**ATG**GCGGCTGGGCTGTTTGGTTTGAGC  
 GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCGCGTCCGCTGGGA  
 ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC  
 CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGTATGAGAAGAACCCA  
 GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT  
 CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA  
 GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC  
 CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**  
**ACC**AGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCTGCCTGCCATTCTGAC  
 CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

235/330

**FIGURE 235**

MAAGLFGLSARLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE  
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER  
LVKYREANGLPIMESNCFDPSKIQLPEDE

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 236**

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC  
TCCCGGCCGCCCGCGTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT  
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA  
CGAAAAC TTGTATGAGAAGAACCAGACTCCCATGGTTATGACAAGGACCCGTTTTGGACG  
TCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC  
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT  
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAC TGCTTCGACCCCAGCA  
AGATCCAG

Downloaded from www.elsevier.com





**FIGURE 238**

MPLALLVLLLLGPGGWCLAEPPrDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY  
RLFpkALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK  
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK  
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSIWELRQTL SVVFDAFITG  
QGKKDWSLFRMFsRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT  
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH  
PYRAFPVLLLDTPVWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS  
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVdWEESPLFNSLFPVSD  
GSNYFVRLYTEPLLvNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA  
KRLANLIRRARGVPPL

**FIGURE 239**

CAAC**ATG**GGGTCCAGCAGCTTCTTGCTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG  
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC  
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG  
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAAGTGAAGAAGGAGGAA  
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC  
TCCTCCTCTACCAGGTGTCCTCAGAA**TGA**TGCTGGGTCCTTTCTACCTCTGGGGGTCACTC  
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC  
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAG  
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT  
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC  
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG  
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT  
TTGCTGTTTATAATCCAA

240/330

## **FIGURE 240**

MGSSSFVLVLMVSLVLTLLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC  
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

**Signal sequence:**

amino acids 1-19

**N-myristoylation sites:**

amino acids 23-29, 27-33, 32-38, 102-108

**WAP-type 'four-disulfide core' domain signature:**

amino acids 49-63

Downloaded from www.jstor.org

**FIGURE 241**

AAACCTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG  
 AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC  
 TCTAGAACCCGACCCACCACC**ATG**AGGTCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG  
 CGTCCAGTGGTTCCTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTA  
 TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT  
 CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA  
 TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCA  
 CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC  
 ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC  
 ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC  
 AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG  
 GTGTGAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA  
 GCACAGAATGCTGGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA  
 CAGCAGTCATCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCTGCCCTTTCCAG  
 AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG  
 GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG  
 TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCAACCTCACTCTC  
 TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC  
 CTTTGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG  
 TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT  
 GCCGTGGTGGGCAACGGGGGCATCCTGAACAACTCCACATGGGCCAGGAGATAGACAGTCA  
 CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC  
 GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT  
 CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCAC  
 CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT  
 TCTGGTTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG  
 TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA  
 TGGTGGCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC  
 AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT  
 CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA  
 GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC  
 CCGGAACTGCCAAAGCCAAGA**ACTGA**ACCGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC  
 CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCTATGGCTCAGACTAA  
 GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT  
 GGCAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG  
 GATTTTAAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCTGCTGAAAAACACTCTT  
 CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG  
 GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC  
 AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCTTG  
 TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACCTGATAATAATAACAAATGATTGTT  
 GTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAA

**FIGURE 242**

MRSCLWRCRHLSQGVQWSLLLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP  
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW  
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG  
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN  
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVVIKASKSLWLQKLFLPNLTLFLDSRHF  
NQSEWDRLEHFAPPPFGFMELNYSLVQKVVTFRPPVPQQQLLLASLPAGSLRCITCAVVGNGG  
ILNNSHMGQEIDSHDYVFRSLGALIKGYEQDVGTRTSFYGFATFSLTQSLILGNRGFKNVP  
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLHLPDFL  
RYMKNRFLRSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW  
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

**Cytoplasmic Domain:**

amino acids 1-10

**Type II Transmembrane Domain:**

amino acids 11-35

**Lumenal catalytic Domain:**

amino acids 36-600

**Ribonucleotide Reductase small subunit Signature:**

amino acids 481-496

**N-glycosylation Sites:**

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

**FIGURE 243**

CGATGCGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG  
GAGCAGCGAGTGGAATTGTTCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT  
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA  
CCAGAGACTGCACCATTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT  
GTCAGTCCAGTGTGCATGGAGGATAAGTGAGCAGACCGTACAGGAGCAGCACACCAGGAGCC  
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG  
ACAAGAGATTTATTTTGCAGACAGACTCTTCATAAGTCCTTTGAGTTTTGTATGTTGTTG  
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

244/330

**FIGURE 244**

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKH FHRLYHNT

RDCTIPAYYKRCARLLTRLAVSPVCMEDK

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
2	2	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
3	3	2	1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
4	4	3	2	1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
5	5	4	3	2	1	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81																			



**FIGURE 245**

GGGCTGGGCCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG  
CCCGACCCCGGCCGCGCCCAGCCCCACC**ATG**CCACCCGCGGGGCTCCGCCGGGCCGCGCCG  
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT  
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT  
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG  
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT  
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCCTGTTGCTACCTGT  
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC  
CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCCAAAGCTGGCCCTGCACCCCCACAGCC  
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC  
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC  
**TGA**GGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT  
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA  
GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA  
ACTATGAGGGGTTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG  
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG  
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT  
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG  
CTAGATTAAAGCTGTAAAGACAAAA

**FIGURE 246**

MPPAGLRRAAPLTAIALLLVLGAPLVLAGEDCLWYLDNRNGSWHPGFNCEFFTFCGTCYHRYC  
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQQLQSP  
FEGQEIPMTGIPVQPVPYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP  
YMPPQPSYPGA

**Transmembrane Domains:**

amino acids 10-28, 85-110

**N-glycosylation Site:**

amino acids 38-41

**N-myristoylation Sites:**

amino acids 5-10, 88-93





**FIGURE 249**

CAGCCACAGACGGGT**CATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTTCATCCTCCCAC  
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC  
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA  
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG  
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC  
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT  
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG  
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT  
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC  
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA  
ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC  
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT  
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG  
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG  
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG  
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCCTGTCCCAGGAGACCGGCAGTGTC  
CTACCTGTGTGCAGCCCCCTTGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG  
GGCGCCACTCATTGTTATGATGGGTACATTTCATCTCTCAGGAGGTGGGCTGTCCACCAAAAT  
GAGCATTACAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG  
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT  
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG  
GTGGGGAGTGGTTTGCCCTTCCTG**C**TAA****CTCTATTACCCCCACGATTCTTCACCGCTGCTGA  
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTTC  
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTATATCTACTCACCTAACAGCA  
ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG  
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

**FIGURE 250**

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI  
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP  
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRRLRGGGIFSNLRVQGCMPPQGCN  
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL  
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN  
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC  
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC  
PSC

**FIGURE 251**

GCGACGGGCAGGACGCCCCGTTGCGCTAGCGCGTGCTCAGGAGTTGGTGTCTGCCTGCGCT  
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG  
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG  
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGACGGCCTGGAAGAGTCG  
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCTGCAT  
GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC  
TGGTCTTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTGCGCAAGGCCATCGGGG  
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCGC  
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA  
CGCCCAGCTGTCTTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA  
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC  
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA  
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT  
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG  
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTTGGGGGCCCCACATGTCCCTGCAGGGTT  
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG  
GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG  
AAAGTGTTCCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA  
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC  
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

**FIGURE 252**

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG  
PTGEKGDMDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRKAIGE  
MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN  
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS  
GGWNDVACHTTMYFMCEFDKENM

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.332660>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



**FIGURE 253**

AGTGACTGCAGCCTTCCTAGATCCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG  
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC  
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACCATGAGGATCATGCTGCTATTACAGCCAT  
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG  
TTCCTGGCGGGGGCCGAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC  
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA  
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA  
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCCTCGGCCC  
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC  
TTTATTAAGACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT  
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC  
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT  
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA  
TAGTGAATATCCCCAACCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT  
CCTACATTAAAAATATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTTGCATATGAA  
AA

## **FIGURE 254**

MRIMLLFTAILAFSLAQSFHAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK  
ALSQASTDPKESTSPEKRDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK  
SSLGTEEQRPL

**Important features:**

**Signal peptide:**

amino acids 1-18

**Tyrosine kinase phosphorylation site.**

amino acids 36-45

**N-myristoylation site.**

amino acids 33-39, 59-65

**Amidation site.**

amino acids 90-94

**Leucine zipper pattern.**

amino acids 43-65

**Tachykinin family signature.**

amino acids 86-92

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGCTGTGCC  
CGCGCTGTGCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG  
GAGCCCTGCGGAGAGCTCAAGCGCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC  
CCATAGTTGCTGCAGGAGTGAGGCC**ATG**AGCTGCGTCCTGGGTGGTGTGCATCCCCTTGGGGC  
TGCTGTTCCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCAACGTCACCTCTCTTAGAGGAG  
CTGCTCAGCAAATACCAGCACAAAGAGTCTCACTCCCGGTCCGCAGAGCCATCCCCAGGGA  
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT  
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC  
CTGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG  
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTGTGTTTGTGTT  
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAAATCGTCATGCCCTG  
AAACCT**TAG**ACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG  
GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT  
TGCCAGGCTGGTCTTGAACCTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG  
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA  
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC  
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCAGGG  
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC  
CTTAGCCCGTGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGCGTGGTTGCCATGACT  
GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG  
CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT  
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG  
TGGAACCTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG  
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT  
TTTAAACAAACCCACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAAATC  
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG  
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC  
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

256/330

## **FIGURE 256**

MSCVLGGV IPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML  
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR  
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation site.**

amino acids 27-31, 41-45

#### **N-myristoylation site.**

amino acids 126-132, 140-146

#### **Amidation site.**

amino acids 85-89



258/330

**FIGURE 258**

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTS  
VTLHHARSQHHVVCNT

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 259**

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT  
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT  
T**ATG**ATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC  
TATTAATATTTACCATTCGAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGGATCTG  
TGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCTCTCC  
CTCCGATTGTTCTAAA**TAA**TTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG  
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTT  
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA  
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT  
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT  
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC  
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA  
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGA  
ATAAAAGTTCATATCTACCC

260/330

## **FIGURE 260**

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP  
SDCSK

**Important features:**

**Signal peptide:**

amino acids 1-29

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.





**FIGURE 262**

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE  
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL  
QEAEKKIRTLNLASCDNMLMGIKSLKIVKKMMDTHGSMKDAVYNPKVYLLIGSRNNTV  
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLEFFHNQATSNEIIKYNLQKRTVED  
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGTLGVEHSWDT  
PCRSQDAEASFLLCGVLYVVYSTGGQGPHERITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH  
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK



**FIGURE 264**

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK  
QYQIRIRKEKPQQHNFTHRLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ  
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK  
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTAAVSPCGRFVASC  
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV  
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH  
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHKLRASNESTRQRLQQQLTQ  
AQETLKSLGALKK

**Important features:****Signal peptide:**

amino acids 1-25

**N-glycosylation site.**

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

**Beta-transducin family Trp-Asp repeat protein.**

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

**FIGURE 265**

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG  
CAGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCC  
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT  
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT  
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA  
AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC  
CGCTGCTCCATGGACTTGAAGAACATCAATTTTT**TAG**GCGCTTGCCTGGTCTCAGGATACCCA  
CCATCCTTTTTCTGAGCACAGCCTGGATTTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC  
TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG  
GCAGACATACTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG  
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA  
AATGGCAGAAAGGACATTCCCCCTCCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG  
CCCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCTTGGGCACAGGCTCTTGGGT  
GCATTGCTCAGAGTCCCAGGTCTTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG  
ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTCAGACCACAGAC  
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA  
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA  
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT  
TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAATGCAAATGGTGGTGGTTCAA  
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACCTCCTTTCCA  
AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG  
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGGCACTGATTCAGACCAGGGAGG  
CAACTACACACCAACATGCTGGCTTTAGAAATAAAAGCACCAACTGAAAAAA

266/330

## **FIGURE 266**

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCCAISLWLRGLRMCTPLGREGECHP  
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

**Signal peptide:**

amino acids 1-19

**Tyrosine kinase phosphorylation site:**

amino acids 88-95

**N-myristoylation sites:**

amino acids 33-39, 35-41, 46-52

266/330

**FIGURE 267**

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTC  
CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC  
TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG  
CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAA  
GCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT  
CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG  
TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT  
TTCAAAGGAGAAATCTTCTGGATGAAAAGAAAAGTTCTATGGTCCACAAAGGCGGAAGAT  
GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTCCGAGCCTGGAACGGAG  
GCTTCTCTGGAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCTGTGGTGGGATCA  
GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT  
TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA**TGAT**  
TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTCATGGGATGTATT  
GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA  
TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA  
CAAAAATCTGAAAACTAATGAGGATTATTAAGCTAAACCTGGGAAATAGGAGGCTTAAAA  
TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG  
TGAGCAAGTCACTTGAGGTCGGGAGTTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC  
TCTACTAAAAATACAAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCG  
GGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA  
CACCCTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

**FIGURE 268**

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKALEYLEDIDLKTLEKEPR  
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF  
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGLGGVFV  
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK



**FIGURE 269**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG  
GGCCAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA  
AGCCCCCTTCCTCGGGCGCTGCCAACCCGCCACCCAGCCCC**ATG**GCGAACCCCGGGCTGGGGCTG  
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCCGCTGGGGCCGAGCCTGGGGGCAAATACA  
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG  
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG  
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG  
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG  
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATC**TAG**GTCCCCTCTCCTGCATCTGTCTCC  
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG  
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG  
GCTATTCACTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

270/330

**FIGURE 270**

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIV  
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**FIGURE 271**

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT  
TTGGGATTTTAATTTTCAAACACAGCAGAATGACATTTTTTCTGTCACTATTATTATTGTTG  
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT  
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC  
CTACCAAAGCTGTCAAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA  
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG  
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAAATGTATCCCTGTCATA  
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT  
TTATTAATTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA  
TCATATAATTTGATACAAATAAAAGAAAAGTGTTCTCTCCCCTTACAGAATTGACATTTTAA  
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA  
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

272/330

**FIGURE 272**

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK  
GIVKGRNLDSRGLILGAEAWGRGVKKNT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 273**

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT  
 GTGCCTGCTGCACCAGTCAAATACTTCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA  
 TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT  
 ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAAGATTTTTTTTTCAAAAA  
 TGTATCTATATTAATTCTTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG  
 AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA  
 TACACCAAGCAGTTTCACAGAATGTGGAGAGAAAGGCCAATACATTTCACTTCACCCCTGACCT  
 TCTACTTGGAaaaaaaacaaaatgaatatggaccaccaggcaaaactgTTTTGTCCATGAGTGGG  
 CTCACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG  
 TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA  
 GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG  
 GAAAAGATTGTCAATTCTTTCTTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG  
 CAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAAAAAACCCATAATCAAGAAGCTCCAAG  
 CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTTCTGAGGATT  
 TTAaaaaacaccataccccatggTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC  
 AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT  
 AAATCGAATGAATCAAGCAGCAAAACATTTCTTGCTGCAGACTGTTGAAAAATGGATCCTGGG  
 TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAAGC  
 AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT  
 CTGCTCTGGAATTAAATATGCATTTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT  
 CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTTGATGAAGTG  
 AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT  
 AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG  
 GCCTCATTTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAAGTCCCTT  
 CAGCTCGAAAGTAAGGGATTAACTACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT  
 TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA  
 TTTCTCTCTGGGATCCCAGTGGAACAATAATGGAAAAATTCACAGTGGATGCAACTTCCAAA  
 ATGGCCTATCTCAGTATTCAGGAACCTGCAAGGTGGGCACTTGGGCATACAATCTTCAAGC  
 CAAAGCGAACCAGAAACATTAACATAAGTAACTTCTCGAGCAGCAAATTTCTTCTGTGC  
 CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCCAATGATT  
 GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTCAT  
 TGAATCACAGAATGGACATACAGAAGTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT  
 CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT  
 AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT  
 GAATAGAGCCGCGTACATAACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA  
 GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTTCAGCCGAACAGCATCCGGA  
 GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA  
 AATCACAGACCTTGATGCCACAGTTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG  
 GAGATAAATTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT  
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGA  
 GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC  
 ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT  
 GCACAAGTAACCTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC  
 TACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT  
 TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTT**TGA**ACCTTA  
 ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTTAAAAAACAAAACAAATGTAAGT  
 AAAGGATATTTCTGAATCTTAAAATTCATCCCATGTGTGATCATAAACTCATAAAAAATAATT  
 TTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATATGTAAAAACTGT  
 CAAGATTAAAAATTTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC  
 AAAGATCCTTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT  
 GATATTTCAAATTTGCATCAAGAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAA  
 GGAGAGCAAATAAACAAACATTTGGAAAAAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa  
 AA

**FIGURE 274**

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTTASTY  
 LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC  
 GEKGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR  
 CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVE  
 FCNEKTHNQEAPSLQNIKCNRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLV  
 LDKSGSMGGKDRLNRMNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM  
 AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGEGDNTASSCIDEVKQSGAIVH  
 FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT  
 LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG  
 TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG  
 YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNQDVYSRYFTAYTENGYSKLVRAHG  
 GANTARLKLRLPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQV  
 PSLPLPDQYPPSQITDLATVHEDKIILTWTAPGDNFQVGVQRYIIRISASILDLRDSFDD  
 ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSANIAQVTLFIP  
 QANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900



MMVVRRLGLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQALALPRANSPTGKEGYQAVLQ  
EWEEQHRNYVSSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL  
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIES  
ALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP  
MKVKNEKLNMAN TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK  
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR  
LNTQPGKKVFYPVLFSQYNPGIIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI  
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPV RGLFHLWHEKRCMDELTP EQYKMCMQS  
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT







**FIGURE 279**

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG  
GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**  
ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG  
CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC  
TCGGTTAAAATGCACCTTCTCCAGCTTTGCCCTGTGGGTGATGCTCTAACAGTGACCTGGA  
ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC  
CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA  
TGCCTCCATCCTTCTCTGGAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA  
AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA  
CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT  
AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC  
ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT  
GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA  
GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT  
TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC  
AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA  
GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAA  
GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC  
AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA  
ATGTGTTTACTCTCTTTCCTTCCACATTCTCAATTAAAGGTGAGCTAAGCCTCCTCGGTG  
TTTCTGATTAAACAGTAAATCCTAAATTCAAAGTTAAATGACATTTTTATTTTTATGTCTC  
TCCTTAACTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATT  
TTTGTCTG

280/330

**FIGURE 280**

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT  
WNFRPLDGGPEQFVFYHIDPFQPMSGFRKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ  
VKNPPDVGVI GEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVL FQHYRKKRWAER  
AHKVVEIKSKEEERLNQEKKVSVYLEDTD

400 420 440 460 480 500 520 540 560 580 600 620 640 660 680 700 720 740 760 780 800 820 840 860 880 900 920 940 960 980 1000

**FIGURE 281**

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACCATGAAGTTCTTAGCAGTCCTGGT  
ACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTG  
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACTGCTGCT  
GCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG  
TAAAGACATTCCAGTTTTACCCAAATGGGTGGGGATCTCCCGAATGGTAGAGTGTGTCCCTT  
GAGATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACCTATTCATGCTTCCTGTGATTTC  
ATCCAACCTACTTACCTTGCCTACGATATCCCCTTTATCTCTAATCAGTTTATTTTCTTTCAA  
ATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

282/330

**FIGURE 282**

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTAT  
TAASTTARKDIPVLPKWVGDL PNGRVCP

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

**FIGURE 283**

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC  
CTAGGGGTCTGGATTTGCTGGTTAACAAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**  
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA  
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC  
ACTTGGAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG  
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG  
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT  
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAA**AACCACAGGCTGG  
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA  
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTTACCACATTAGCAATTAAACTGAGAAAT  
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT  
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA  
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG  
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT  
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

284/330

**FIGURE 284**

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHLQPFVTERTLGKVQRWSGVHTQTGGRAG  
GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

284/330



**FIGURE 285**

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA  
GCGGCCCCCATGGGCGGCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG  
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA  
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC  
CGGGGCCGGGATGCAGCCCAGGAACCTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA  
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA  
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT  
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC  
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC  
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAAGTCTGAG  
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG  
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC  
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA  
CCCTTTCATGCCTACACACCCCTCATTAAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA



**FIGURE 287**

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAAATTTGCAT  
CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTA AAAAGCCAGGA  
AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC  
TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC  
TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG  
AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC  
CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA  
CATGGTTCACGGAAGGCAAGTTTGTGTGACGTCAACGGAATCGCTATCTCCTTCTCAACTGGG  
ACCGTGACAGCCTAACGGTGGCAAGCGAGAAAACCTGTGTCTTCTCCCAATCAGCTCAG  
GGCAAGTGGAGTGATGAGGCCTGTGCGAGCAGCAAGAGATACATATGCGAGTTCACCATCCC  
TAAATTAGGTCTTTCTCCAATGTGTCTCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA  
TCTCTAAGATCAAGTAAAAATCATAATTTTTTACTTATTAAAAAATTGCAACACAAGATCAAT  
GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT  
TCTTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT  
AAACAGACTAAAATCTTTCTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTTTTCAA  
AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT  
TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT  
GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC  
TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT  
GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCC  
TACCCTTTTTTTTGGAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT  
CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAAACCTAAGGCATAT  
CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT  
AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTTTTTACATTT  
TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT  
TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGGA AAAATTGCATACGTCTGTGCAATT  
TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT  
TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAGCAAACAATTTT  
AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT  
CCCACTTTGCAAACCTTAACTACACATGCTTGGAATTAAGTTTTTAGCTGTTTTCATTGCTCA  
ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSHKRRVRDKDGLKTQIEKLWT  
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ  
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSSQAQ GK  
WSDEACRSSKRYICEFTIPK

**FIGURE 289**

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC  
CCCGAGCCCCCGCGCC**ATGA**AGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA  
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCAGCCTGTCGCTGCGCTG  
GAGTCGGCGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCCTCGGCACCCTCAACCCGCT  
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT  
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG  
GCCCTGACAGTGTTTGGC**TGA**GCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC  
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCGGCCCCCT  
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

290/330

**FIGURE 290**

MKLAALLGLCVALSCSSAAAFVGSAPVAPVAALESAAEAGAGTLANPLGTNLNPLKLLLS  
SLGIPVNHLEGSQKCVNELGPQAVGAVKALKALLGALTVFG

290/330

TGAAGGACTTTTCCAGGACCCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACCTCTTGGCCTCCGCAGCCGATCAC**ATGA**AGGTGGTGCCAAGTCTCCTGCTCTCCGTCTCCTGCGACAGGTGTGGCTGGTACCCGGCTTGGCCCCAGTCTTCAGTCGCCAGAGACCCCAGCCCCTCAGAACCAGACCAGCAGGGTAGTGCAGGCTCCAGGGAGGAAGAGGAAGATGAGCAGGAGGCCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTTGCCAAGGAGACTTCAAACCTTCGGATTACAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGGCAACATGGTCTTCTCTCCATTTGGCATGTCTTGGCCATGACAGGCTTGATGCTGGGGGCCACAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAGCCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAAGTGGGCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCTTCAATTTATCCAAGAGGTATTTTGATACAGAGTGCCTGCCTATGAATTTTCGCAATGCCTCACAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACGTTTTGATGAGATTAATCCTGAAACCAAATTAATTTCTTGTTGATTACATCTTGTTCAAAGGGAATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCACCTGGACAAGTACAAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAATTTTCTGTTGTATGTCTCAAACCTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCTCATGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACATGGCTCAGAAACATGAAAACCAGAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCAGAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTGCTGACCTTAGTGAACCTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGACAGTGATTGAAGTTGATGAAAGGGGCAGTGAAGCAGTGGCAGGAATCTTGTCAGAAATTACTGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTTCATTTTCATGATCTATGAAGAAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TAA**ATTCAGACATGCATAAGCACTTCGTGCTGTAGTAGATGCTGAATCTGAGGTATCAAACACACACAGATACCAGCAATGGATGGCAGGGGAGAGTGTTCTTTTGTCTTAAGTATGTTTAGGGTGTCTCAAATAAATACAGTAGTCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGCTGAAACGGTGGACAGTGCTGAACCTTATATATATTTTTTCTTACACATACATACCTATGATAAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAAATGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACCTGATTATAGAGAAGGCTACTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCAATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGCTGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTTCATTTAATGTTTTTGGACCATGGTGGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAAATGATTAATTTGATACATATTTTTTTAAAAA

**FIGURE 292**

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E  
E E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I  
K R G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F  
D T E C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P  
V F T E V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L  
A L E D Y L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A  
T G R N L Q V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F  
L G R V V N P T L L

099974 12004  
Total = 126860



**FIGURE 293**

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**CATG**AG  
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG  
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG  
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTGCTGTTCCC  
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTGAGGGCAGGGGCCCCA  
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC  
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC  
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA  
TCTACCACCCCCAG**TAG**GGGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGG  
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCAGCAGGCAAA  
AAAAAAAAAAAAAAAAA

294/330

**FIGURE 294**

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL  
FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE  
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 295**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTG  
 TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAACTCAGCTTCCTGC  
 TGTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA  
 TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC  
 TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCT  
 GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG  
 CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC  
 AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG  
 ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG  
 CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC  
 TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT  
 ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC  
 GACGCCCAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATT  
 TGTTCAAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG  
 TCACCGGATGTAACACTGAGCATCACTGCATTTGGTGGAGGAGGATACTTTCCAGAGGCCAGT  
 CCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATATGGAACCATGTTGGTTA  
 CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG  
 GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACCTACCCA  
 GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

296/330

**FIGURE 296**

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN  
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG  
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI  
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYYSPTYGQREFTAGFVQFRVFNNERAAN  
ALCAGMRVTGCNTEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLL  
FYR

296/330

**FIGURE 297**

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC  
CACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGCGAGGTGCTTGGGCCG  
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC  
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC  
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC  
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG  
GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC  
TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG  
TAACCCACAATAGTTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT  
TCTGAAGCAAAGAAAGGATCAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC  
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC  
GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA  
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAACAATATTCT  
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA  
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTGAATAAACATCTGGATCTTATAGACCGT  
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCT  
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA  
TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT  
AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA  
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAACATCAACATTTTTTG  
TGCTTTAACTGTAGTAGTTGGTCTAGAAACAAATACTCC



**FIGURE 299**

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCAGCCGGGAGCCGG  
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCAGCG**ATG**GCGACCCTGTGGGGAGGC  
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTCGTGCCTGGCGCTTTCCGTGCTGCTGCTGGC  
GCAGCTGTCAGACGCCGCCAAGAATTTTCGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT  
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT  
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA  
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA  
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG  
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT  
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG  
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG  
CATGTTGTCCTCAGC**TAA**TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA  
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTACCAACT  
GTTGCTGGAAGATTCAAAACCTGGAAGCAAAAACCTTGCTTGATTTTTTTTTTTCTTGTTAACGTA  
ATAATAGAGACATTTTTTAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG  
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA  
AGCACTCTCTTTTTTACCACATAGTTTTTAACTTGACTTTCAGATAATTTTTCAGGGTTTTTG  
TTGTTGTTGTTTTTTGTTTGTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGGTT  
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT  
TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG  
ACTTTTGCCTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT  
CTAAATGCCTGGTGGCTTTTTCACAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG  
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG  
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT  
TGCAATAAAGAAATTTTATTTTAAACCAAGCCTCCCTGGATTGATAATATATACACATTTG  
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT  
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT  
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTGAAACATCAA





GCACCTGCGACCACCGTGAGCAGTCAATGGCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT  
CTGGGCTTGTCCTGGCTCTGTGCTGCTGCTGCCAAGGCCTTCCTGTCCCGCGGGAAGCGG  
CAGGAGCCGCCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA  
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG  
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG  
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTTATATACTGTACATTCTATT  
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAATTGAAAATCTAATATGGCGATAAAAA  
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA  
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAGTTTCACATAAGAATG  
TTTACTCAATGTTTAAGTGTTTTGCCCCAAAATTCACAATAACAAGGCAGAACTAGGACTT  
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

302/330

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPTPEGKLGRFPMMHHHQAPSDGQT  
PGARFQRSHLAEAFAKAKSGGGAGGGSGRGLMGQIIPIYGFGLFLYILYILFKVSRIILI  
ILHQ

[illegible]

**FIGURE 303**

CGGCTCGAGTGCAGCTGTGGGGAGATTTAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT  
GGATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC  
TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTTCCCGCCTGAGCTAACAGTCCATGTG  
GGTGATTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTTCAAGAT  
AGACTGGACTCTGTACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA  
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC  
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA  
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG  
AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAGATGGGATGTGTTTTCCAG  
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA  
GGAGATTGTATTTCTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG  
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCATCATG  
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT  
GGTGTTCAGAAAACCATTTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC  
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC  
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA  
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG  
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA  
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA  
CCCAGTTTGGCCTTCTCTGAGGTGAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG  
GAATGCCAAAAACACAGCAAGCCTTT**TGAGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG  
TGGAGACTCTCTCCTGTGTGTGTCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC  
CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACTGAAGATGGAGAATTTGGAGCCTGG  
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC  
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTT  
GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA  
AAAACCAACCCAAATCAA

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG  
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES  
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY  
HKLRMSVEYSQSWGHFQNRVNLVGDI FRNDGSIMLQGVRESGGNYTCSIHLGNLVFKKTIV  
LHVSPEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTTCGNKSSVNSTV  
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR  
SDRNN SLEKKSGGGMPKTQQAF

**FIGURE 305**

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG  
GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG  
AAACTACATTTTGC AAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA  
GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG  
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG  
ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA  
TGAGAATGAAAATCGCACAGGAAGCTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG  
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC  
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTAACATGGGAAGAGAG  
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG  
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTTCGGATTATCTCGCCAGAAGTCGAAT  
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGGAAGA  
TGGAAAAGGAAATATGAATTGTGCTTATTTTTCATAATGGGAAAATGCACCCTACCTTCTGTG  
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT  
**TAA**TGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT  
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

306/330

**FIGURE 306**

MQDEDGYITLNIKTRKPALVSVGPASSSSWVRVMALILLILCVGMVVGLVALGIWSVMQRNYL  
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE  
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL  
EDGKGNMNCAYFHNGKMHPTFCENKHLYLMCERKAGMTKVDQLP

306/330

**FIGURE 307**

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG  
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC  
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA  
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG  
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAAGAAGCGGTGGTGGTGGG  
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG  
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTGAGCAGCCCCAGCAAAGGCAAGACCAGCTGC  
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG  
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC  
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT  
CTGTTTAAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT  
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA  
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG  
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA  
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT  
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC  
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC  
**G****TAG**CCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT  
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG  
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

**FIGURE 308**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFSGSKRRRRRP  
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA  
MNSEGILYTSELFTECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLENKEGEIMKGNHVK  
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKRSVSGVLNGGKSMHNEST

**N-glycosylation site.**

amino acids 242-246

**Glycosaminoglycan attachment site.**

amino acids 165-169, 218-222

**Tyrosine kinase phosphorylation site.**

amino acids 93-100

**N-myristoylation site.**

amino acids 87-93, 231-237

**ATP/GTP-binding site motif A (P-loop).**

amino acids 231-239

**HBGF/FGF family proteins**

amino acids 78-94, 102-153



**FIGURE 309**

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG  
ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG  
CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT  
CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTTCCCTTCCTGACATTGGCAGTG  
CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC  
AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA  
ACCACGTCTTGAAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG  
GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAACTCGAGGGTGGGATCCACTGAGG  
AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT  
GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG  
CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG  
CGCGCTCCGGGCGCCTGCCGTTTGGGGGTGTCTCCTCCCGGGGCGCTATGGCGGCGCTGGC  
CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTGCG  
CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG  
CTGTCCAAGGTGCGACTGTGCGGGGGGCGGCCCCGCGCGGCCGACCGCGGCCCCGAGCCTCA  
GCTCAAAGGCATCGTCACCAAACCTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG  
ACGGAAGCATCCAGGGCACCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT  
GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC  
TGAGGGACTGCTCTACAGTTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT  
TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCC  
TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA  
GGCAGCTGCCCACCTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC  
ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCCTGAAATGTAGTCCCTGGACTG  
GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

310/330

**FIGURE 310**

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR  
GPEPQLKGIVTKLFCRQGFYLLQANPDGSIQGTPEDTSSFTHENLIPVGLRVVTIQSAKLGHY  
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR  
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

**Tyrosine kinase phosphorylation site:**

amino acids 199-207

**N-myristoylation sites:**

amino acids 54-60, 89-95, 131-137

**HBGF/FGF family signature:**

amino acids 131-155

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 311**

**ATG**GCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG  
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG  
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG  
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT  
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT  
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT  
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT  
TAAAGAATCTGTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG  
AATCTGGTAGAGCCTGGTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA  
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG  
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA  
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

**FIGURE 312**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWD RPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR  
RQDPQLKGIVTRLYCRQGYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY  
IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR  
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKFPVNKSKTT

**N-glycosylation site.**

amino acids 100-104, 242-246

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 28-32, 29-33

**Tyrosine kinase phosphorylation site.**

amino acids 199-207

**N-myristoylation site.**

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

**HBGF/FGF family proteins.**

amino acids 104-155, 171-198

**FIGURE 313**

[illegible]

**FIGURE 314**

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC  
 VRKNLREVPDGI STNTRLLNLHENQIQI IKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA  
 NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS  
 YISEGAFEGLSNRLRYLNAMCNLREIPNLTPLIKLDLDELDSGNHLSAIRPGSFQGLMHLQKL  
 WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNDIL  
 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE  
 LKCRASLTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTNLNFTNVTQDTGMYTCMVNSVGN  
 TTASATLNVTAATTTTFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ  
 STRSTEKTFTIPVTDINS GIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN  
 HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS  
 VHEPLLIRMNSKDNVQETQI

**Signal sequence:**

amino acids 1-44

**Transmembrane domain:**

amino acids 523-543

**N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,  
 434-438, 442-446, 488-492, 606-610

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

**Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

**N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,  
 391-397, 422-428, 433-439, 531-537

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCCGCACAT  
GGCTGCAGCCACCTCGCGCGCACCCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTGCGGA  
GGCGCCCCGGCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCCGCGTCCGGGGATC  
GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA  
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC  
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCATAATGAAGGGAACCAAAAA  
GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCC  
AGTGGCCTTTTGCTTCCAATTTCTTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC  
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT  
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC  
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCATTGTGTATT  
ACTGGCAGCGAATCCGAGAGAAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT  
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA  
CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAAGTGTACAGT  
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG  
ATTTTCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA  
GAGACCTAATGAAATTGAGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT  
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCACTCGCTCCACAGCAAAT  
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCCCAGCCAGGGCTGGCCAC  
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG  
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA  
ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC  
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA  
GTGAGCATTGCACGGAACAGATTGAGATGAGCATTTTCTTATACAATACCAAACAAGCAAA  
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG  
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG  
AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTTGTATCAGTGCCTTGATTACAAAT  
TTCAAGAGGAAATGGGATGCTGTTTGTAATTTTCTATGCATTTCTGCAAATTTATTGGATT  
ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC  
TGAGCTAACCACCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC  
TTCATTTGTGATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA  
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC  
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCATCTTCATGATGTT  
ATGAGGATTGTTGACAAACATTTAGAAATATATAATGGAGCAATTGTGGATTTCCTTCAAT  
CAGATGCCTCTAAGGACTTTCTGTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT  
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA  
CCCAACATACCATTTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG  
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA  
TGGTGCCAGGCACCTGTAGGAAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC  
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

## **FIGURE 316**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYVGTGLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV  
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV  
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID  
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI  
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS  
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 232-251

316/330



**FIGURE 317**

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCTACCTGCAGCCGCCGCCACGGCACGGCAGCCA  
 CCATGCGCCTCCTGCTGTGCTTCCGTGCTCCTGTGCGGAGTAGTGGATTTGCGCCAGAAGTTTGAGTATCACTACT  
 CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA  
 CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT  
 ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC  
 AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGTGAGATATTGGCACATATCAGTGCAAAGTGAA  
 AAAAGCTCCTGGTGTGCAAATAAGAAGATTCTCTGGTAGTTCTTGTAAAGCCTTCAGGTGCGAGATGTTACG  
 TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG  
 TATGAGTGGCAAAAATTTGCTGACTCACAGAAAATGCCACCTTCATGGTTAGCAGAAATGACTTCATCTGTTAT  
 ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC  
 AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT  
 TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGCTGTGCTAAAAAGCGCAGAGAAGAAAAATATGAAAA  
 GGAAGTTCATCAGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG  
 GCAGTAATCATTATCCCTGGGGTCCATGCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA  
 GTACCAAGTGAAGACTTTGAACGCACCTCCTCAGAGTCCGACTCTCCACCTGCTAAGTTCAAGTACCCTTACAA  
 GACTGATGGAATTACAGTTGTAATAAATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT  
 AGGCCTCTAGTAAAGACTTAAATGTTTTTAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC  
 ATCTACTTTTATGCAATGGCATTAGACATGTAAGTCAGATGTGATGTCAAATTTAGTACGAGCCAAATTTCTTGT  
 TAAAAAACCTATGTATAGTGACACTGATAGTTAAAAAGATGTTTTATATATTTTCAATAACTACCACTAACAA  
 ATTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTCAA  
 AGGAAATTTTAAATTTCTTACGTTCTGTTAATGTTTTTGTCTATTTAGTTAAATACATTGAAGGGAAATACCCG  
 TCTTTTTCCCTTTTATGCACACAACAGAAACGCGCTGTGTCATGCCTCAAACATTTTTTTTATTTGCAACTACA  
 TGATTTACACAATTTCTTAAACAACGACATAAAATAGATTTCTTGTATATAAATAACTTACATACGCTCCA  
 TAAAGTAAATTTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTCTCGTATCCAACAGAGTTGATGC  
 ACAATATATAAATACTCAAGTCCAATATTA AAAA ACTTAGGCACTTGACTAACTTTAATAAAATTTCTCAAAC TA  
 TATCAATATCTAAAGTGCATATATTTTTTAAGAAAGATTATTCTCAATAACTTCTATAAAAAATAAGTTTGATGG  
 TTTGGCCCATCTAACTTCACTACTATTAGTAAGA ACTTTTAACTTTTAAATGTGTAGTAAGGTTTATTCTACCTT  
 TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT  
 TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTCTTCCC  
 TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCCTGGCTT  
 TTTTCTGGGCAAAGGTTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT  
 TAAGTTAATTTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAGACAAAGTTACCCACAACACT  
 GAGATCACATCTAAGTGACACTCCTATTGTGAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA  
 TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCTAAACAA  
 CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAAGGACATGCTTGTTTTAGTCCAGTGGTTT  
 CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT  
 AGCAAAC TAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT  
 TATGAACACAAAGAAGGAAACAAATAGACATTGGAGTCTATTTGAGAGGGGAGGTTGGGAGAAGGAAAAGGAGCA  
 GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA  
 TGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAAATAAAAGTTAAAAA AAAAAAAAAAAAAA  
 AA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 318**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA  
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK  
APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT  
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSQCLLRNLNVPPSNKAGLIAGAIIGTLL  
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM  
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

**Signal sequence.**

amino acids 1-19

**Transmembrane domain:**

amino acids 236-257

**N-glycosylation sites.**

amino acids 106-110, 201-205, 298-302

**Tyrosine kinase phosphorylation sites.**

amino acids 31-39, 78-85, 262-270

**N-myristoylation sites.**

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,  
245-251, 296-302

**Myelin P0 protein.**

amino acids 96-125

**FIGURE 319**

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA  
ATGGTGAAGGTGCCTGTCTAACTTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC  
CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTTCGC  
CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG  
AGAGATGAATTTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC  
CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG  
CTTCTCTGGAGCCAGGTATCAGGGGGCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT  
GAAGGGGGTGTTCCTCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG  
CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT  
GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTTTTTCAAAAA  
CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC  
TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC  
AGTGCACACAGGCGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC  
TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC  
TCTGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCATTTCA  
AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT  
TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG  
GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTT  
AATTAATGTCAGTATTTCAACTGAAGTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG  
CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG  
TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTGGCCTGTCTTTGGATT  
GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACCTGAACTTCAGAGCATGAAAATCACACT  
GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAG  
TAAACGATAAAATGTGGATTAAAGTGCCAGCACAAAGCAGATCCTCAATAAACATTTTCATT  
TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTT  
TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC  
ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA  
CTCCAAAAA

320/330

## **FIGURE 320**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMVMVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK  
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVFKNH  
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL  
TKALGEVDILLTWMQKFYKL

### **Signal sequence:**

amino acids 1-42

### **N-glycosylation sites.**

amino acids 85-89, 99-103, 126-130

**FIGURE 321**

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC  
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC  
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG  
ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA  
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTC AAGGATCA  
TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA  
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC  
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA  
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT  
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT  
GTGGGAGACAGCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC  
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGC AAAAGTCTACTGTGGTAT  
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT  
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT  
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCATATTTTACCTATGA

**FIGURE 322**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST  
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ  
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

**Signal sequence:**

amino acids 1-18

**N-glycosylation sites.**

amino acids 56-60, 135-139

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 102-106

**N-myristoylation site.**

amino acids 24-30

**Actinin-type actin-binding domain signature 1.**

amino acids 159-169

**FIGURE 323**

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG  
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT  
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC  
GGTTCTATCGATAATCTCAGCACCAAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGGCCCGCCT  
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA  
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC  
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC  
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA  
TGAGCAGAAGATACCTCTGCATGGATTTTCTAGAGGCAACATTTTTGGATCACACTATTTTCGAC  
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC  
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC  
CACCCCCGTACTCCAGTTCTGTCCCGGAGGAACGAGATCCCCCTAATCACTTCAACACC  
CCCATAACACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCTGAACGT  
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTACAGGAGCTCCCGAGCG  
CCGAGGACAACAGCCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTTCGAGTGAAC  
ACGCACGCTGGGGGAACGGGCCCCGAAGGCTGCCGCCCTTCGCCAAGTTCATC**TAG**GGTTCG  
CTGG

**FIGURE 324**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARNSYHLQIHKNGHVD  
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLNGY  
DVIHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE  
RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA  
KFI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 175-179

**N-myristoylation site.**

amino acids 33-39, 100-106, 225-231, 229-235

**HBGF/FGF family proteins**

amino acids 73-124





**FIGURE 326**

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL  
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSNFMVLIATSVETSAAASGSPEGAGM  
TTVQTITGSDPEEAI FDTLCTDDSSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSD  
GPHPVITPSRASESSASSD  
GPHPVITPSRASESSASSD  
GPHPVITPSWSPGSDVTLLAEALVTVTNI  
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL  
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY  
VKVSGAAPVSI EAGSAVGKTTSFAGSSASSYS  
PSEALKNFPTSETPTMDIATKGPFP  
TSRDLPLSPVPTTTNSSRGTNSTLAKITTSAKTTMKPOQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

### Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,  
246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,  
404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

GCGGAGCATCCGCTGCGGTCTCTCGCGAGACACCCCGCGGATTTCGCCGGTCTTCCCGCGG  
GCGGAGACAGAGCTGTCTCGCAGCCTGGATTGGCAGCAGGGGCGCGGGGTCTCTCGACGCCA  
GAGAGAAATTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT  
CCTTGACCTTTGAAGACCAAACTAAACTGAAATTTAAAATGTTCTTCGGGGGAGAAGGGAG  
CTTGACTTACACTTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT  
GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTCTAAGGGGAATC  
AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAA  
AAAAACATATCAGGGGACAAAGCATGTAACCTGATGATCTTCGACACTCGAAAAACAGCTA  
GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA  
AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG  
CCAAGAGTTACCCAGGAAGATTCTCTCTTACATGGCCAATTTTACAAGCAGTCACTCCCC  
TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT  
CAGAAGTTTGGATCCTCAGATCACCTGGAGAAACTATTTAAGATGGATGAAGCAAGTGCCCA  
GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTTACAATTTTCTCTGATCAAGAAA  
TAGCTCATCTGCTGCCTGAAAATGTGAGTGCCTCCGACTACGGTGGCAGTTGCTTCTCCA  
CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC  
TTCTGGGACTTCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC  
AGCCTCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG  
GCTACAACAGCAGTTCTGACTACCACCTTTGAGGCACCTACGGACTCGAAAGGCAGCTTAGA  
AACCATAACGTTTACAGAAATCTCCAACCTTAACTTTGAACACAGGGAATGTGTATAACCCTA  
CTGCACTTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCCTGGGAAGGT  
AGGGAGGCCAGTCCAGGCAGTTCCTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC  
ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTGTTTCTGGTGATAGGCC  
TCGTCTCTCTGGGTAGAATCCTTTGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT  
TATTTGATCAATGGGATCTATGTGGACATCTAAAGGATGGAACCTCGGTGTCTCTTAATTCAT  
TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG  
TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTTCTTTTTTTTTTTGGAGACAGAGTCTT  
GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC  
CTGGGTTCAAAGCATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA  
CCACACCTGGGTGATTTTTGTATTTTTTAGTAGAGACGGGGTTTACCATGTTGGTCAGGCTG  
GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG  
CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTTGAGAAGGAATGAAGTG  
GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT  
ATGTAAAGTAATAAAGTATAATTGCCATATAAAATTTCAAATTCAACTGGCTTTTTATGCAA  
GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC  
TGTTTTATATCAATTTCTAATGGATTTGCTTTTTCTTTTTATATGGATTCCTTTAAACTTAT  
CCAGATGTAGTTCCCTTCCAATTAAATATTTGAATAAATCTTTTGTACTCAA

**FIGURE 328**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED  
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP  
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF  
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL  
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP  
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV  
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRDLYLINGIYVDI

**Signal sequence.**

amino acids 1-25

**Transmembrane domain.**

amino acids 384-405

**N-glycosylation sites.**

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 415-419

**Tyrosine kinase phosphorylation site.**

amino acids 50-57

**N-myristoylation sites.**

amino acids 4-10, 48-54, 315-321

**FIGURE 329**

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCTGCTATGGGGTTGCCTGCTGCT  
CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT  
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT  
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT  
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA  
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG  
TCTTTACTGATCTCTCTGTTCTTTCCAGGACCCTGCTGTCTCCTCCCCTCCCCTTCTCCCAC  
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC  
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG  
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC  
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC  
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG  
GTGTCCATCCCAGATGGTCCGCATACTGGCCCCAGTCTTGGTGTGCTGAGCCTTCTGTGAGC  
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA  
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC  
CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA  
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT  
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCAGAAAGCTTTCCACCTCAGCCTCAGAG  
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGCATGTTCCA  
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG  
GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA  
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGC  
CTCATGCCAGTGTGCGACCCCTGCCTTCTCCCCTCCAGACCCACCTTGTCTTCCCTCCC  
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGTGATGAAGAGGAGCATGCT  
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT  
GTGAAAAACGTGATTCTTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG  
GACTCTGAATTCTAACAATGCCAGTGAAGTGTGCGCACTTGAGTTTGGGGCCAGTGGGCCTG  
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC  
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG  
TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG  
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA  
AAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTCTTGGTGGCTAGAGA  
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG  
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA  
GCACAACTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT  
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCTCCTGGGTTCAAGTGATT  
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT  
TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCCTGAC  
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG  
TCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTGTTGAGTATGCAAACTTGGAAAG  
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT  
TATTTGTTTTTGTGTTGTTACTTCCCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT  
TTTACAGAGCAATTATCTTGTATATACAACCTTGTATCCTGCCTTTTCCACCTTATCGTTCC  
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA  
GCTGCATAAAAAAAAAAAAAA

**FIGURE 330**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPTEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS  
GTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLTLDAGEYWCGVEKRGPD ELLISLFV  
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG  
TSQYGHERTSQYTGTSPHPATSPPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI  
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD  
VISMPPLHTSEEEELGFSKFVSA

**Important features:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 248-269

**N-glycosylation site.**

amino acids 96-99

**Fibrinogen beta and gamma chains C-terminal domain.**

amino acids 104-113

**Ig like V-type domain:**

amino acids 13-128